

534,419
Rec'd PCT/PTO 06 MAY 2005

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
27 May 2004 (27.05.2004)

PCT

(10) International Publication Number
WO 2004/043999 A2

(51) International Patent Classification⁷: C07K 14/47

IN7 (CA). SAMUELS, Mark [US/CA]; 2743 West 37th Avenue, Vancouver, British Columbia V6N 2T5 (CA).

(21) International Application Number:
PCT/CA2003/001716

(74) Agent: MBM & CO.; P.O. Box 809, Station B, Ottawa, Ontario K1P 5P9 (CA).

(22) International Filing Date:
12 November 2003 (12.11.2003)

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/425,601 12 November 2002 (12.11.2002) US
60/502,453 12 September 2003 (12.09.2003) US

(84) Designated States (*regional*): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): LAFRENIERE, Ronald, G. [CA/CA]; 1264 Osborne Avenue, Verdun, Québec H4H 1X5 (CA). DUBÉ, Marie-Pierre [CA/CA]; 4831 Hutchison Street, Montreal, Québec H2V 4A4 (CA). MACDONALD, Marcia, L [CA/CA]; 102-1345 West 13th Avenue, Vancouver, British Columbia V6H

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HSAN II RELATED GENE AND EXPRESSION PRODUCTS AND USES THEREOF

(57) Abstract: Methods of utilizing a gene related to pain perception, herein dubbed HSN2, or its encoded protein, dubbed herein "sensorin," for the screening and identification of agents for the treatment of pain, neuropathy and related disorders, especially small organic compounds, as well as methods of using these compounds to treat or otherwise ameliorate pain, neuropathy and related disorders in human patients. Novel polypeptides and polynucleotides, along with their nucleotide and amino acid sequences, are also disclosed.

BEST AVAILABLE COPY

WO 2004/043999 A2

HSAN II RELATED GENE AND EXPRESSION PRODUCTS AND USES THEREOF

5 This application claims priority of U.S. Provisional Applications
60/502,453, filed 12 September 2003, and 60/425,601, filed 12 November
2002, the disclosures of which are hereby incorporated by reference in their
entirety.

FIELD OF THE INVENTION

10
15 The present invention relates to methods of utilizing a gene (HSN2)
involved in Hereditary Sensory and Autonomic Neuropathy, type II (HSAN II)
and expression products of this gene for the screening and identification of
agents, such as small organic compounds, useful in the treatment of pain,
neuropathy and related disorders in human patients as well as uses thereof to
20 treat or otherwise ameliorate such disorders.

BACKGROUND OF THE INVENTION

25
30 The present invention relates to the medical disorder called Hereditary
Sensory and Autonomic Neuropathy – Type II ("HSAN II"), which is a member
of a group of hereditary pain disorders known as the Hereditary Sensory and
Autonomic Neuropathies (HSAN). HSAN comprises a group of five clinically
and genetically heterogeneous disorders which are mainly characterized by
variable sensory and autonomic dysfunction, including absence of pain. The
pathology of HSAN is characterized by degeneration of peripheral sensory
35 neurons and it appears to be hereditary, making it of considerable scientific
interest to determine which gene or genes underlies the disorder.

HSAN II is an autosomal recessive condition with distal generalized sensory loss. (See Online Mendelian Inheritance in Man (*OMIM*) reference *201300) Patients with HSAN II typically present with "glove and sock" distribution, meaning that they have reduced or complete loss of pain, temperature and touch sensations in the lower legs and feet as well as hands and forearms. Age of onset is quite early, usually in the first decade of life. They first complain of a numbness in their extremities which is aggravated by the cold, after which pain sensation is reduced. Patients have difficulty with handling small objects (like coins) in their pockets due to the lack of sensation. Some families appear to have a slowly progressive form, where the lack of sensation slowly grows up the legs and arms until in some cases, the patients have reduced sensations in their trunkal regions while others do not. Reduced sensation is confined just to the lower legs (usually below the knees), forearms and hands. Muscle atrophy (particularly in calves, and in between the thumb and first finger), ulcerations and infections are major secondary features. Amputations of fingers, toes and sometimes feet and lower limbs, resulting from infections, are not uncommon. Reflexes are diminished or lost. There is very little, if any, autonomic involvement, so sweating and blood pressure are normal.

20

The pathology of HSAN II appears to be tied to peripheral nerve degeneration. Biopsy reveals that within the sural nerve there is a severe loss of myelinated axons, and a lesser loss of nonmyelinated fibres. No cutaneous sensory receptors or nerve fibers are seen. (Axelrod, F.B. "Autonomic and sensory disorders". Chapter 117; pages 3146-3161 in Emery and Rimoin's Principles and Practice of Medical Genetics (Fourth Edition) Edited by: David L. Rimoin, J. Michael Connor, Reed E. Pyeritz, Bruce R. Korf. Published: Harcourt Publishers Limited London 2002).

25

The genetic basis of HSAN II is yet to be determined. No genetic loci have been confirmed, and no mutations in suggested candidate genes have been identified.

30

The HSAN II gene (dubbed herein "HSN2") would present a novel target for therapeutic agents. The functional validation disclosed herein shows that the activity of the gene/protein is involved in very specific human disease processes. Therapeutic agents which modulate the biological activity of the HSAN II gene or its corresponding protein are novel therapeutic agents for the treatment of pain, neuropathy and related disorders.

The screening assays disclosed herein select from a large library of compounds those compounds that interact with, bind to or otherwise modulate the activity of the target gene/protein. A wide body of commercial literature describes the screening of chemical libraries of diverse potential therapeutic agents to identify such potential therapeutic agents.

Therapeutic agents for treatment of pain fall into two main classes - the NSAIDs (non-steroidal anti-inflammatory drugs) and the opioids. NSAIDs treat pain in a way similar to the mechanism of aspirin, the most well-known and oldest member of the class. Common NSAIDs include acetaminophen, ibuprofen and naproxen. These drugs mainly inhibit the body's ability to synthesize prostaglandins. The common mechanism of action for all NSAIDs is the inhibition of the enzyme cyclooxygenase (COX). A major commercial success has been achieved with specific inhibitors of COX-2, such as Celebrex™ from Pharmacia/Pfizer, and Vioxx™ from Merck & Co.

Opioids act through the opioid receptor family. These drugs include the weak opioids such as codeine and Tylenol 3, and strong opioids such as morphine and methadone. Some are long acting, others are of short duration. Opioid analgesics have a tendency to addiction and dependency, and so are not preferred for long-term or chronic pain management.

Outside of the NSAIDs and opioids, there are a number of other suggested analgesic agents in clinical trials (i.e. not yet approved for marketing) which are believed to have alternative targets. Some clinical trials are attempting to establish that central neuropathic pain may respond to ion channel blockers such as blockers of calcium, sodium and/or NMDA (N-

methy-D-aspartate) channels. The literature provides substantial pre-clinical electrophysiological evidence in support of the use of NMDA antagonists in the treatment of neuropathic pain. Such agents also may find use in the control of pain after tolerance to opioid analgesia occurs, particularly in cancer patients.

Current treatment of neuropathies first entails identifying the cause, then treating the cause if possible. There is a substantial list of disorders that can result in neuropathies, including diabetes and other diseases. Most neuropathies (except the genetic forms) are treatable by treating the root causes, though most treatments are only partially effective. For example vitamin deficiency is treated with vitamins; Guillain-Barre and CIDP are treated with gamma globulins and plasmapheresis. For genetic forms of neuropathy, current clinical practice is to manage patients by supporting their efforts to avoid self-inflicted injury.

In accordance with the present invention, the identification of the hereditary basis for HSAN II facilitates developing more potent agents for treating pain, neuropathy, and related disorders. The underlying genetic mutation provides a therapeutic target for novel therapeutic agents. This therapeutic target permits identification and discovery of more effective analgesics and other therapeutic agents, as well as new methods and compositions for diagnosis of HSAN II distinguish between types of inherited pain disorders.

25

BRIEF SUMMARY OF THE INVENTION

In one aspect, the invention relates to the nucleic acid sequence for HSN2, including the genomic sequence, mRNA or cDNA, polymorphic, allelic, isoforms (adult, neo-natal, etc.) and mutant forms thereof, and nucleic acid constructs of the gene, including vectors, plasmids and recombinant cells and transgenic organisms containing or corresponding to HSN2 (or knock-outs thereof).

In another aspect, the invention relates to the gene product of HSN2 (sensorin), including the polypeptide, protein, and amino acid sequence, and the polymorphic, allelic, isoforms (adult, neo-natal, etc.) and mutant forms thereof, and recombinant cells and transgenic organisms wherein this polypeptide or a polypeptide corresponding thereto is expressed.

In another aspect of the present invention, the HSN2 gene or protein is incorporated into a screening assay whereby test compounds (potential therapeutic agents) are evaluated for their ability to modulate HSN2 gene expression or sensorin activity, thereby identifying modulators of the gene or protein and thus potential therapeutic agents.

In preferred embodiments, the polynucleotide whose expression is to be measured or monitored is present in an intact cell, preferably a mammalian cell, most preferably a peripheral neuron, and may include a recombinant cell. Such polynucleotide may also be present outside of a cell and the expression may be measured *in vitro*.

In another aspect, the present invention relates to a method for identifying an agent that modulates the activity of a polypeptide encoded by a polynucleotide as disclosed herein.

In a further aspect, the present invention relates to a method for identifying an agent for the treatment of pain, neuropathy or a related disorder in an animal. Preferably, the animal is a mammal, such as a human being. In

specific embodiments, the pain stimulus is a heat stimulus and reaction or sensitivity to hot and/or cold may be measured. In another embodiment, an electrical stimulus may be used:

- 5 In a preferred embodiment of the invention, the compound identified which modulates HSN2 (or sensorin) is selective for the target ahead of related genes or proteins.

10 In a further aspect, the present invention relates to a method for treating a condition in an animal afflicted with chronic pain, neuropathy or a related disorder comprising administering to said animal an effective amount of an agent first identified by an assay method of the invention. Preferably, said animal is a human patient. Such determined agents may also be applied to alternative or additional indications beyond pain, neuropathy or other disorders, which are found to be treatable by modulating HSN2 or sensorin
15 activity.

 In a further aspect, the present invention relates to compounds that modulate HSN2 or sensorin. These compounds include antibodies, antisense compounds, gene therapy vectors and proteins, and small molecule organic compounds. The invention also comprises the use of these compounds in the
20 treatment of pain, neuropathy and related disorders.

BRIEF DESCRIPTION OF THE DRAWINGS

25

 Figure 1a shows the pedigree for a family designated HSAN4 showing family relationships and haplotypes. Figure 1b shows the pedigree for a family designated HSAN3 showing family relationships and haplotypes. Squares and circles represent males and females, respectively. Filled
30 symbols indicate individuals with HSAN II. Symbols with an "N" indicate individuals diagnosed as normal, and clear symbols and symbols with a question mark indicate individuals who have not been diagnosed

Figure 2. Domain prediction of the HSN2 protein, sensorin. TM topology was predicted using TMPred.

Figure 3. Peptide sequence alignment of human, mouse, rat, fugu and
5 tetraodon sensorin proteins.

Figure 4. Peptide sequence alignment of human and Fugu sensorin.

Figure 5. Peptide sequence alignment of Fugu and Zebrafish orthologs
10 of sensorin.

Figures 6A-6E shows the sequences of HSN2 from various sources. In all cases, cDNA sequences are shown in capital letters. Start and stop codons are underlined.

15

Figure 6A shows a sequence (SEQ ID NO: 1) for wild-type human HSN2 putative cDNA (with ends based on human ESTs).

Figures 6B-1 and 6B-2 shows the promoter sequence (SEQ ID NO: 6)
20 for human HSN2 (presumably there only 1 promoter used and it is not associated with a CpG-rich island; this corresponds to the segment of PRKWNK1 intron 8 upstream of the HSN2 start codon; the transcription start site and cDNA sequence (shown in capital letters) for HSN2 is tentative and based on 1 EST sequence; the translation start site, and start ATG, is the last
25 3 bases.

Figure 6C shows the nucleotide sequence (SEQ ID NO: 7) of wild-type mouse HSN2 cDNA predicted (5' and 3' ends not determined). The putative polypeptide is SEQ ID NO: 8.

30

Figure 6D shows the nucleotide sequence (SEQ ID NO: 9) for wild-type rat HSN2 (with the genomic sequence defined by 5' and 3' ends in EST BF522762/AI578184, and containing a polyA tail).

Figure 6E-1 to 6E-6 show the genomic structure and sequence of the human HSN2. (SEQ ID NO. 12), with the figures following in sequence. The HSN2 gene is a single unspliced exon found in intron 8 of the PRKWINK1 gene (also called XH03, below). Here, italics = interspersed repeats, uppercase = exons, bold = exons (or part of exons) found in common isoforms and underlined = start and stop codons, or upstream in-frame stop codon.

Figures 7A-7C show the sequences of cDNAs for three separate mutations identified in individuals afflicted with HSN II. Start and stop codons are underlined. Figure 7A shows mutation 1 (SEQ ID NO: 14) as human cDNA with mutation c.594delA. Figure 7B shows mutation 2 (SEQ ID NO: 15) as human cDNA with mutation c.918insA. Figure 7C shows mutation 3 (SEQ ID NO: 16) as human cDNA with mutation c.943C>T.

Figure 8 shows three in-frame candidate translation initiation codons.

Figure 9 shows a sequence alignment between the human and pig peptide sequences.

DEFINITIONS

The following terms have their indicated meaning unless expressly stated otherwise herein.

As used herein, the term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

$$\text{Percent Identity} = 100 [1-(C/R)]$$

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence for which the percent identity as calculated above is about equal to or greater than a specified minimum Percent Identity then the Compared Sequence has the specified minimum percent identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

As used herein, the terms "portion," "segment," and "fragment," when used in relation to polypeptides, refer to a continuous sequence of nucleotide residues, sequence forms a subset of a larger sequence. Such terms include the products produced by treatment of said polynucleotides with any of the common endonucleases, or any stretch of polynucleotides that could be synthetically synthesized. These may include exonic and intronic sequences of the corresponding genes.

30

As used herein the term "encode" and its derivatives mean "can be transcribed into" or "can be translated into" so that a gene or other deoxy polynucleotide sequence can be described in terms of its ability to encode an RNA or polypeptide while the latter terms relate to materials that can be

described in terms of their being encoded by said gene or deoxy polynucleotide. In the same way, an RNA can be said to encode a polypeptide while said polypeptide can be described as being encoded by said RNA.

5 The words "therapeutic target" mean a gene, an RNA or polypeptide for which therapeutic intervention is achieved with agents that modulate the activity of said gene, RNA or protein. "Modulate" means to increase, to decrease, or to otherwise change the expression or activity, especially of a therapeutic target.

10

 "HSN2 activity" or "sensorin activity" as used herein, especially relating to screening assays, is to be interpreted broadly and contemplates all directly or indirectly measurable and identifiable biological activities of the HSN2 gene, gene products and sensorin. The HSN2 gene is thus involved with
15 functions related to Hereditary Sensory and Autonomic Neuropathy - Type II ("HSAN II") and is, in the past, has also been called the "HSAN II Susceptibility Gene."

 The term "polynucleotide" is used interchangeably with "gene", "cDNA",
20 "mRNA", "oligonucleotide", and "nucleic acid".

 The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue
25 is an analog or mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

 The term "agent" is used interchangeably with the term "compound" and likewise the term "test agent" is used interchangeably with the term "test
30 compound".

 The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence,

wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

The term "mammal" refers to a member of the zoological class
5 *Mammalia*. Examples of mammals include, without limitation, humans, primates, chimpanzees, mice, rats, rabbits, sheep, and cows.

The term "treat" or "treatment" encompasses therapeutic treatment, preventive treatment and protective treatment. Such protective treatment
10 includes treatment of patients, especially to achieve analgesia.

As used herein, the term "correspond" applied to a gene or polynucleotide, such as a DNA, means that the gene has the indicated nucleotide sequence or that it encodes substantially the same RNA as would
15 be encoded by the indicated sequence, the term "substantially" meaning at least 95% identical as defined elsewhere herein.

The term "DNA segment" refers to a DNA polymer, in the form of a separate fragment or as a component of a larger DNA construct, which has
20 been derived from DNA isolated at least once in substantially pure form, i.e., free of contaminating endogenous materials and in a quantity or concentration enabling identification, manipulation, and recovery of the segment and its component nucleotide sequences by standard biochemical methods, for example, using a cloning vector. Such segments are provided in the form of
25 an open reading frame uninterrupted by internal nontranslated sequences, or introns, which are typically present in eukaryotic genes. Sequences of non-translated DNA may be present downstream from the open reading frame, where the same do not interfere with manipulation or expression of the coding regions.

30

"Isolated" in the context of the present invention with respect to polypeptides (or polynucleotides) means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a

living organism is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the co-existing materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polynucleotides, and recombinant or immunogenic polypeptides, disclosed in accordance with the present invention may also be in "purified" form. The term "purified" does not require absolute purity; rather, it is intended as a relative definition, and can include preparations that are highly purified or preparations that are only partially purified, as those terms are understood by those of skill in the relevant art. For example, individual clones isolated from a cDNA library have been conventionally purified to electrophoretic homogeneity. Purification of starting material or natural material to at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. Furthermore, claimed polypeptide which has a purity of preferably 0.001%, or at least 0.01% or 0.1%; and even desirably 1% by weight or greater is expressly contemplated.

The term "coding region" refers to that portion of a gene which either naturally or normally codes for the expression product of that gene in its natural genomic environment, i.e., the region coding *in vivo* for the native expression product of the gene. The coding region can be from a normal, mutated or altered gene, or can even be from a DNA sequence, or gene, wholly synthesized in the laboratory using methods well known to those of skill in the art of DNA synthesis.

30

In accordance with the present invention, the term "nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the proteins provided by this invention are assembled from cDNA fragments and short oligonucleotide linkers, or from a series of

oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon.

5 The term "expression product" means that polypeptide or protein that is the natural translation product of the gene and any nucleic acid sequence coding equivalents resulting from genetic code degeneracy and thus coding for the same amino acid(s).

10 The term "fragment," when referring to a coding sequence, means a portion of DNA comprising less than the complete coding region whose expression product retains essentially the same biological function or activity as the expression product of the complete coding region.

15 The term "primer" means a short nucleic acid sequence that is paired with one strand of DNA and provides a free 3'OH end at which a DNA polymerase starts synthesis of a deoxyribonucleotide chain.

20 The term "promoter" means a region of DNA involved in binding of RNA polymerase to initiate transcription.

25 The term "open reading frame (ORF)" means a series of triplets coding for amino acids without any termination codons and is a sequence (potentially) translatable into protein.

30 As used herein, reference to a DNA sequence includes both single stranded and double stranded DNA. Thus, the specific sequence, unless the context indicates otherwise, refers to the single strand DNA of such sequence, the duplex of such sequence with its complement (double stranded DNA) and the complement of such sequence.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, the human HSN2 Gene, is disclosed. This gene, when mutated, results in Hereditary Sensory and Autonomic Neuropathy - Type II ("HSAN II"). This gene is designated herein
5 as HSN2 (but in the past has sometimes been referred to as SENS or XH18 are other designations). The polypeptide is referred to as HSN2 Protein or HSN2 peptide. In the past, it has also been referred to as HSAN II polypeptide, HSN2 polypeptide, HSANII peptide. It is also referred to as
10 sensorin, although this latter term has been applied to a minor unrelated protein in the past. This naming is not essential to use of the invention claimed herein.

Thus, the present invention provides a gene and its corresponding
15 protein related to insensitivity to pain, neuropathy and related disorders, referred to as HSN2 and its corresponding polypeptide sensorin or HSN2 protein.

The wild type nucleic acid sequence of human HSN2 is provided at
20 SEQ ID NO. 1. The wild type amino acid sequence of human HSN2 protein is set forth in SEQ ID NO: 2.

Three in-frame candidate translation initiation codons have been identified (Fig. 8). The 3rd ATG is closest to the optimal Kozak consensus C-
25 C-[AG]-C-C-A-T-G-G context for initiation of translation, but is not conserved in the rat sequence. Those skilled in the art are capable of characterizing the most efficient translation initiation context by generating three HSN2 cDNA expression vectors, each containing different starting site combinations including one, two or three ATG codons.

30

The human HSN2 gene maps to the HSAN II candidate interval at 12p13.33. The gene encodes a full-length conserved peptide of 434 amino acids. It comprises a single unspliced exon which maps to intron 8 of the PRKWINK1 gene and is transcribed from the same strand as PRKWINK1.

HSN2-encoded protein may be the full length PRKWINK1 protein with a special exon. Thus, the present invention can also encompass a full length PRKWINK1 protein (including cDNA) containing the HSN2 exon.

5 In one aspect, the present invention relates to an isolated polynucleotide comprising a polynucleotide having a nucleotide sequence with at least 60% identity to, preferably at least 70% identity to, more preferably at least 78% identity to, even more preferably at least 90% identity to, most preferably at least 95% identity to, and especially at least 98% identity to a
10 sequence selected from the group consisting of SEQ ID NO: 1, 6, 7, 9 and 12 wherein said isolated polynucleotide encodes a polypeptide that binds to an antibody specific for a polypeptide having the amino acid sequence of SEQ ID NO: 2. In an especially preferred embodiment, the isolated polynucleotide comprises the nucleotide sequence of SEQ ID NO: 1, 6, 7, 9 or 12.

15

In another aspect, the present invention relates to an isolated polypeptide comprising a polypeptide having an amino acid sequence with at least 90% identity to, preferably at least 95% identity to, more preferably at least 98% identity to and most preferably having the amino acid sequence
20 selected from the group consisting of SEQ ID NO: 2, 8, 10, 11 and 13 wherein said polypeptide binds to an antibody specific for a polypeptide having the amino acid sequence of SEQ ID NO: 2.

As disclosed herein, three separate mutations were identified in
25 individuals affected with HSN II, none of which are found in the homozygous state in unaffected individuals (see Table 3, SEQ ID NO: 3 to 5). Each mutation affects the open reading frame (ORF) of the HSN2 gene. The cDNA sequences for these mutations are shown in Figure 7 and may be summarized as follows:

30

Mutation 1. c.594delA; frameshift in codon 198, protein truncated to 206 aa: A deletion of an A (adenine) at position 594 in the nucleotide sequence (relative to the start ATG), found homozygous in the HSN4-70 sample, causes a frameshift in codon 198 leading to premature truncation of

a 206 aa peptide. The predicted amino acid sequence of the truncated human sensorin sequence from Mutation 1 is set forth at SEQ ID NO. 3. The same mutation was also later found to be responsible for the disease in the HSN3 family. This deletion segregated with disease in other members of the HSN3 and HSN4 families. The mutation was not detected in any of 217 normal chromosomes from individuals of European descent.

Two different mutations were found in the HSN5-302 sample.

10 Mutation 2. c.918insA, frameshift in codon 307, protein truncated to 318 aa. This mutation is a 1 bp insertion of an A between bases 918-919, causing a frameshift in codon 307, and leading to premature truncation to produce a 318 aa (amino acid) peptide. The predicted amino acid sequence of the truncated human sensorin sequence from Mutation 2 is set forth at SEQ
15 ID NO. 4

 Mutation 3. c.943C>T; Gln315stop; protein truncated to 314 aa. This mutation is a C>T change at position 943 that changes codon 315 (CAG, encoding glutamine) to a TAG stop codon, therefore prematurely truncating
20 the protein at 314 aa. Each of these mutations was not found in the control sample, dbSNP database, nor available genomic and EST sequences. The predicted amino acid sequence of the truncated human sensorin sequence from Mutation 3 is set forth at SEQ ID NO. 5. The same mutation was also later found to be responsible for the disease in the HSN6 family.

25

 Segregation of mutations in HSN2 with affected HSN II patients has now been confirmed in 5/5 affecteds and 0 control (unaffected) individuals.

30 Promoter: A sequence containing the promoter sequence is set forth at SEQ ID NO. 6. It includes the portion of PRKWNK1 intron 8 upstream of the HSN2 translation initiation site. There are a number of DNA elements that are conserved between human and mouse in the region upstream of the presumed transcription initiation site that may act as promoter elements, including silencers and enhancers, and other transcription factor binding sites.

Other single nucleotide polymorphisms (SNPs) and variations: there is only 1 documented SNP (rs1054186) in the HSN2 gene, which corresponds to a C>T change at position 1690 of the cDNA, located in the 3'UTR. This putative SNP may not be real, as it is found in only one published expressed sequence tag (EST) and the quality of the sequencing is too poor to evaluate additional SNPs in this gene.

BAC contig: The HSN2 gene maps to BAC (Bacterial Artificial Chromosome) clone AC004765, and there are no gaps in the genomic structure. Based on the structure of the HSN2 gene determined above, the gene spans a maximum of 3 kb, consists of a single exon mapping within intron 08 of the XH03 gene (PRKWINK1), and is transcribed from the same strand as XH03, namely tel 5'-HSN2-3' cen in 12p13.33. There are probably no additional spliced exons, since these would interfere with splicing of XH03. The position of HSN2 within an intron of the PRKWINK1 gene is conserved in human, mouse, rat and fish (for example, Fugu and Tetraodon).

Markers: D12S91 is the closest marker, mapping ~29 kb telomeric to HSN2.

Genomic sequence: the full human genomic sequence illustrating the nucleic acid sequence around HSN2 is set forth at SEQ ID NO. 12.

Conservation of HSN2 in other organisms.

Orthologs of the human HSN2 gene were found in mouse, rat and fish (Takifugu rubripes ("Fugu"), Danio rerio ("Zebrafish"), and Tetraodon nigroviridis ("Tetraodon").)

A conserved mouse ORF, and flanking sequences were identified from a mouse BAC AC113092. The nucleic acid sequence of mouse HSN2 is set forth at SEQ ID NO. 7. The translated amino acid sequence of mouse HSN2

is set forth at SEQ ID NO. 8. No mouse ESTs of HSN2 were identified in a comprehensive public database search.

Likewise the rat ortholog was identified from rat genomic sequences, and verified with the sequence of one rat EST (BF522762). This EST was used to identify the conserved rat ORF based on the genomic sequence found in BAC sequences AC106348 and AC106932. The nucleic acid sequence of wild-type rat HSN2 is set forth at SEQ ID NO. 9. The translated amino acid sequence of wild-type rat sensorin is set forth at SEQ ID NO. 10.

Analysis of human, mouse and rat orthologs. The human protein is 87% identical to mouse, and 85% identical to rat. The upstream stop codon immediately preceding the start ATG is conserved in human, mouse and rat.

Fugu (*Takifugu rubripes*): The human sensorin peptide sequence was used to screen the fugu genome by sequence similarity searching. One hit gave an E value score of $1e-25$ against a Fugu genomic fragment (CAAB01000768). The sequence was extracted and the conserved ORF is set out at SEQ ID NO. 11.

Zebrafish (*Danio rerio*): The Fugu peptide sequence was used to screen the non-human non-mouse EST database using tBLASTn to identify additional orthologs. An EST from a zebrafish adult retina cDNA library was found (BG304539) which gave a partial protein sequence which is well-conserved to sensorin in Fugu. There were no additional zebrafish ESTs in GenBank.

Tetraodon (*Tetraodon nigroviridis*): The Fugu peptide sequence was used to screen the publicly available Tetraodon genomic scaffolds (<http://fugu.hgmp.mrc.ac.uk/blast/>) and a partial sequence of the Tetraodon sensorin ortholog was identified. (SEQ ID No. 13)

Alignment of the human, mouse, rat, Fugu, Tetraodon and zebrafish sensorin amino acid sequences is shown in Fig 3.

Evaluation of Sensorin Function

The membrane topology of the sensorin was assessed using TMPred (located at www.ch.embnet.org/software/TMPRED_form.html), which
5 predicted a single N-terminal trans-membrane TM domain, with the N-terminus intracellular. TMPred was used to assess the membrane topology of the sensorin protein. The strongly preferred model predicted only 1 transmembrane domain (score 533, considered significant if >500) located between aa 4 and 22, with the N-terminus located inside the cell. The result of this search is set
10 out at Figure 2.

SignalP, an established signal peptide prediction tool, indicated that the sensorin protein has a signal peptide that is cleaved after the first 17 amino acids. Sensorin may be compact and globular in shape.

15

Further aspects of sensorin function at the biochemical level may be determined by those skilled in the art using standard techniques. It is clear from the instant invention that sensorin is a survival factor for peripheral nerves, as the absence of this gene in humans leads to the neuropathic
20 condition of HSAN II.

Expression: Based on ESTs, the HSN2 gene is expressed in the sympathetic trunk, kidney, testis, fetal liver spleen, head-neck tumor, germ cell tumors, muscle, uterus, retina, whole eye, and CNS (multiple sclerosis
25 lesions). SAGE expression was detected in pancreatic adenocarcinoma, brain astrocytoma, brain glioblastoma, colon adenocarcinoma, ependymoma and normal vascular endothelial cells. There is no GNF data. Therefore the HSN2 gene appears to be widely expressed in many tissues, but at extremely low levels. HSN2 may be expressed in peripheral sensory neurons or
30 supporting Schwann cells, which usually make up only a small percentage of any given tissue's mass.

It is recognized that those skilled in the art may prefer to use forms of HSN2 or sensorin corresponding to the sequences disclosed herein, although

not necessarily the same. For example, screening assays may utilize HSN2 or sensorin from a different organism, preferably a vertebrate, and most preferably from a mammalian species. The shared technical features of these forms of HSN2 or sensorin, are that, when expressed, they have similar biological activity, and that they share functional similarity with HSN2 or sensorin, as the case may be, such as may be determined by those skilled in the art. Thus the invention encompasses the use of, including but not limited to, sheep, dog, rat, mouse or horse HSN2 or sensorin, for the same purposes as set out more specifically herein for human HSN2 or sensorin. The HSN2 gene and/or sensorin polypeptide according to the invention may also be obtained from other mammalian species, other vertebrates, invertebrates and microorganisms based on the disclosure herein.

Because of the processing that may take place in transforming the initial RNA transcript into the final mRNA, the sequences disclosed herein may represent less than the full genomic sequence. They may also represent sequences derived from alternate splicing of exons. Consequently, the genes present in the cell (and representing the genomic sequences) and the sequences disclosed herein, which are mostly cDNA sequences, may be identical or may be such that the cDNAs contain less than the full genomic sequence. Such genes and cDNA sequences are still considered corresponding sequences because they both encode similar RNA sequences. Thus, by way of non-limiting example only, a gene that encodes an RNA transcript, which is then processed into a shorter mRNA, is deemed to encode both such RNAs and therefore encodes an RNA corresponding to an HSN2 sequence as disclosed herein. Thus, the sequences disclosed herein correspond to genes contained in the cells and are used to determine relative levels of expression because they represent the same sequences or are complementary to RNAs encoded by these genes. Such genes also include different alleles and splice variants that may occur in the cells used in the processes of the invention.

Method of Treatment Using HSN2 as a Therapeutic Target.

The discovery that mutations in HSN2 relate to clearly definable physiological outcomes in humans (namely HSAN II) establishes a clear function for this protein, and now allows the inventors to establish, for the first time, that HSN2 Gene and sensorin are useful as therapeutic targets in humans for the treatment of pain, neuropathy and related disorders. Standard industrial processes are available to those skilled in the art to confirm the identity of the therapeutic agents which modulate the activity of the gene or protein, many of which are set out below.

In accordance with the foregoing, the present invention relates to a method for treating a disorder comprising administering to a person in need of such treatment an effective amount of a selective sensorin agonist or antagonist, or a pharmaceutically acceptable salt thereof, or a pharmaceutical composition containing either entity.

In a preferred embodiment, said disorder is pain, neuropathy or a related disorder. In a further preferred embodiment, said administering is by oral or intravenous means.

Identification of Therapeutic Agents.

The present invention also readily affords different means for identification of agents for treating pain, neuropathy and related disorders according to their ability to modulate the activity of HSN2 or its protein. Such means involve testing libraries of chemical compounds, either one at a time or in combinations, in an assay format which is designed to measure a biological activity related to HSN2 or sensorin, the protein encoded by HSN2. Those library compounds that modulate the biological activity in the desired fashion are thereby identified as therapeutic agents of the invention.

Exemplary assay methods useful for the identification of such compounds are detailed herein, although those skilled in the art will be aware

of alternative means. In a first step, compounds are sequentially tested against the assay to determine whether they influence a measurable biological activity of the assay.

5 Assays may be based one or more of the diverse measurable biological activities of a gene or polypeptide corresponding to HSN2 or sensorin. Relating to the purified sensorin protein, sensorin activity includes, but is not limited to, all those biological processes, interactions, binding behavior, binding-activity relationships, pKa, pD, enzyme kinetics, stability,
10 and functional assessments of the protein. Relating to sensorin activity in cell fractions, reconstituted cell fractions or whole cells, these activities include, but are not limited the rate at which sensorin performs any measurable biological characteristic and all measurable consequences of these effects, including cell growth, development or behavior and other direct or indirect
15 effects of sensorin activity. Relating to HSN2 genes and transcription, HSN2 activity includes the rate, scale or scope of transcription of genomic DNA to generate RNA; the effect of regulatory proteins on such transcription, the effect of modulators of such regulatory proteins on such transcription; plus the stability and behavior of mRNA transcripts, post-transcription processing,
20 mRNA amounts and turnover, and all measurements of the rate or amount of expression and translation of the mRNA into polypeptide sequences. Relating to HSN2 activity or sensorin activity in organisms, this includes but is not limited biological activities which are identified by their absence or deficiency in disease processes or disorders caused by aberrant HSN2 or sensorin
25 biological activity in those organisms. Broadly speaking, HSN2 biological activity or sensorin biological activity can be determined by all these and other means for analyzing biological properties of proteins and genes that are known in the art.

30 The invention therefore provides numerous assays which measure an activity of HSN2 or sensorin and are useful for the testing of chemical compounds to identify which ones affect such activity.

In one aspect, the present invention relates to a method for identifying an agent that modulates the activity of the HSN2 gene, comprising:

a) contacting a test compound with a cell that expresses an HSN2 gene;

5 b) determining a change in the expression of said gene as a result of said contacting,

wherein said determined change in expression of the gene indicates modulation, thereby identifying the test compound as an agent that modulates the activity of an HSN2 gene.

10

In a preferred embodiment of such method, HSN2 is a mammalian HSN2 gene, most preferably where the mammal is mouse, rat or human. Also preferred is where determined modulation is a decrease in expression and/or where the gene is present in a cell, preferably a mammalian cell, preferably a nervous system cell, such as a neuron. Such cells may include a recombinant cell, especially one genetically engineered to contain and express the HSN2 gene. In a preferred embodiment, said gene has the sequence of SEQ ID NO: 1.

20 In other embodiments, the gene comprises a polynucleotide corresponding to a polynucleotide having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 7 and 9.

In another aspect, the present invention relates to a method for identifying an agent that modulates HSN2 gene, comprising:

25 (a) contacting a test compound with a genetic construct comprising a reporter gene operably linked to an HSN2 promoter under conditions where the reporter gene is expressed;

30 (b) determining a change in expression of the reporter gene as a result of said contacting,

wherein a determined change in expression indicates modulation, thereby identifying the test compound as an agent that modulates HSN2 gene.

In a preferred embodiment thereof, the modulation is a decrease in expression of said reporter gene. Normally the reporter gene is some readily measured gene and is not HSN2 but the HSN2 gene, with promoter sequence, may be used and still lie within the contemplation of the invention, even though the gene is not contained within a cell. Also preferred is where the HSN2 promoter is a mammalian HSN2 promoter, especially where the mammal is mouse, rat or human. Most preferred is where the promoter has the nucleotide sequence of SEQ ID NO: 6 and/or where the genetic construct is present in a cell, preferably a mammalian cell, such as a nervous system cell, preferably a neuron. Of course, such construct will normally be present in a recombinant cell.

The present invention also contemplates a method for identifying an agent that modulates the activity of an HSN2-encoded protein, comprising:

- a) contacting a test compound with an HSN2-encoded polypeptide under conditions where said polypeptide is active; and
- b) determining a change in the activity of said polypeptide as a result of said contacting;

wherein said determined change in activity indicates modulation, thereby identifying the test compound as an agent that modulates the activity of HSN2-encoded protein (such as sensorin).

In a preferred embodiment thereof, the HSN2-encoded polypeptide is a mammalian HSN2-encoded polypeptide, preferably where the mammal is mouse, rat or human. Also preferred is where the determined modulation is a decrease in activity of the polypeptide.

Also preferred is where the polypeptide is present in a mammalian cell, especially where this has been engineered to contain or express said polypeptide, such as by genetic engineering. In one such embodiment, the recombinant cell does not express said polypeptide absent said engineering. Also preferred is where the cell is a cell of the nervous system, especially a neuron, most especially a neuron involved in pain response.

Also preferred is where the polypeptide is encoded by a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1, 7, and 9, especially where the polypeptide comprises an amino acid sequence of
5 SEQ ID NO: 2, 8, 10 or 11.

In a preferred embodiment, the observed change in activity in step (b) is a decrease in activity and is the result of binding to said polypeptide by said chemical agent of step (b). Also preferred is where the agents are useful for
10 treating pain, neuropathy or a related disorder.

In other preferred embodiments, the polypeptide is part of an intact cell, and the present invention specifically contemplates embodiments in which the cell is engineered by other than genetic engineering, such as where the
15 activity of a polypeptide is to be enhanced and the cell has been engineered to contain, or have on its surface, said polypeptide but wherein the polypeptide is present due to physical insertion of the polypeptide into the membrane or cytoplasm of the cell and not through expression of a gene contained in the cell. Methods well known in the art, such as use of
20 polyethylene glycol, viruses, and the like, are available to effect such insertions and the details of such procedures need not be further described herein.

In one preferred embodiment of such method, the polypeptide is a
25 polypeptide that reacts with and/or binds to, an antibody that binds to, or is specific for, a polypeptide having an amino acid sequence at least 95% identical to, more preferably at least 98% identical to, the sequence of SEQ ID NO: 2 and where any difference in amino acid sequence is due only to conservative amino acid substitutions. In an especially preferred embodiment,
30 the polypeptide has the amino acid sequence of SEQ ID NO: 2.

The sensorin assays of the invention may employ compound screening technology such as (but not limited to) the ability of various dyes to change

color in response to changes in assay conditions resulting from the activity of the polypeptides. Compound screening assays can also be based upon the ability of test compounds to modulate the interaction of the target peptide (sensorin) and known interacting proteins. Such interacting proteins can be identified by a variety of methods known in the art, including, for example, radioimmunoprecipitation, co-immunoprecipitation, co-purification, and yeast two-hybrid screening. Such interactions can be further assayed by means including but not limited to fluorescence polarization or scintillation proximity methods.

10

Agents that have the effect of modulate the half-life of sensorin in cells would also act to induce decreased sensitivity to pain and thereby achieve analgesia. By way of non-limiting example, cells expressing a wild-type sensorin are transiently metabolically labeled during translation, contacted with a candidate compound, and the half-life of the polypeptide is determined using standard techniques. Compounds that modulate the half-life of the polypeptide are useful compounds in the present invention.

In one such assay for which the polypeptides encoded by genes disclosed herein are useful, the polypeptide (or a polypeptide fragment thereof or an epitope-tagged form or fragment thereof) is bound to a suitable support (e.g., nitrocellulose or an antibody or a metal agarose column in the case of, for example, a his-tagged form of said polypeptide). Binding to the support is preferably done under conditions that allow proteins associated with the polypeptide to remain associated with it. Such conditions may include use of buffers that minimize interference with protein-protein interactions. If desired, other proteins (e.g., a cell lysate) are added, and allowed time to associate with the polypeptide. The immobilized polypeptide is then washed to remove proteins or other cell constituents that may be non-specifically associated with it the polypeptide or the support. The immobilized polypeptide can then be used for multiple purposes. In a compound screening embodiment, compounds can be tested for their ability to interfere with interactions between sensorin and other bound molecules (which are presumably sensorin

interacting proteins). Compounds which can successfully displace interacting proteins are thereby identified as sensorin modulating agents of the invention.

In an alternative embodiment designed to identify sensorin interacting proteins, the immobilized polypeptide is dissociated from its support, and proteins bound to it are released (for example, by heating), or, alternatively, associated proteins are released from the polypeptide without releasing the latter polypeptide from the support. The released proteins and other cell constituents can be analyzed, for example, by SDS-PAGE gel electrophoresis, Western blotting and detection with specific antibodies, phospho-amino acid analysis, protease digestion, protein sequencing, or isoelectric focusing. Normal and mutant forms of such polypeptide can be employed in these assays to gain additional information about which part of the polypeptide a given factor is binding to. In addition, when incompletely purified polypeptide is employed, comparison of the normal and mutant forms of the protein can be used to help distinguish true binding proteins. Such an assay can be performed using a purified or semipurified protein or other molecule that is known to interact with a polypeptide encoded by an HSN2 polynucleotide.

This assay may include the following steps.

1. Harvest the encoded polypeptide and couple a suitable fluorescent label to it;
2. Label an interacting protein (or other molecule) with a second, different fluorescent label. Use dyes that will produce different quenching patterns when they are in close proximity to each other versus when they are physically separate (i.e., dyes that quench each other when they are close together but fluoresce when they are not in close proximity);
3. Expose the interacting molecule to the immobilized polypeptide in the presence or absence of a compound being tested for its ability to interfere with an interaction between the two; and
4. Collect fluorescent readout data.

An alternative assay for such protein interaction is the Fluorescent Resonance Energy Transfer (FRET) assay. This assay can be performed as follows.

1. Provide the encoded protein or a suitable polypeptide fragment
5 thereof and couple a suitable FRET donor (e.g., nitro-benzoxadiazole (NBD)) to it;
2. Label an interacting protein (or other molecule) with a FRET acceptor (e.g., rhodamine);
3. Expose the acceptor-labeled interacting molecule to the donor-
10 labeled polypeptide in the presence or absence of a compound being tested for its ability to interfere with an interaction between the two; and
4. Measure fluorescence resonance energy transfer.

Quenching and FRET assays are related. Either one can be applied in
15 a given case, depending on which pair of fluorophores is used in the assay.

Additionally, drug screening assays can also be based upon polypeptide functions deduced upon antisense interference with the gene function. Intracellular localization of pain-related polypeptides, or effects
20 which occur upon a change in intracellular localization of such proteins, can also be used as an assay for drug screening.

In accordance with the foregoing, the present invention provides the amino acid sequence of a protein, designated sensorin or HSN2 Polypeptide, that is found in neuronal cells (for example, SEQ.ID NO: 2 from humans) and
25 which is associated with hereditary transmission of insensitivity to pain and neuropathy. In addition, several mutations, have been found in this sequence derived from individuals found to have such indifference to pain. Thus, agents that mimic the phenotypic effects of this mutations, such as aberrant protein structure and decreased, or absent, function represent candidate compounds
30 for evaluation as therapeutic agents of the invention.

Relating to expression assays, in one aspect the present invention relates to a method for identifying an agent that modulates the activity of a polynucleotide whose expression contributes to pain sensation, comprising:

a) contacting under physiological conditions a chemical agent with a polynucleotide corresponding to the promoter region of HSN2 (in a preferred embodiment, the promoter has the sequence of SEQ ID NO: 6); and

b) detecting a change in the expression of said polynucleotide as a
5 result of said contacting;

thereby identifying an agent that modulates said polynucleotide or gene activity.

Such modulation is preferably a decrease or an increase in expression.

10 In preferred embodiments, such expression is measured by measuring the amount of an expression product. In a convenient embodiment, the promoter region of HSN2 is operably linked to a reporter gene, that is, a gene which whose expression is conveniently measured (for example, reporter genes such as Green Fluorescent Protein, luciferase, chloramphenicol acetyl-
15 transferase (CAT), and the like).

In preferred embodiments, the polynucleotide whose expression is to be measured or monitored is present in an intact cell, preferably a mammalian cell, most preferably a neuronal cell. In additional preferred embodiments,
20 such an intact cell is a cell that has been engineered to comprise said polynucleotide, such as by genetic engineering, most preferably wherein the cell does not express the subject gene or polynucleotide absent having been engineered to do so.

25 In accordance with the disclosure herein, upstream untranslated regions and promoter regions of HSN2 are readily obtained from SEQ ID No. 12 and other publicly retrievable sequence databases. Such genomic or untranslated regions may be included in plasmids comprising the identified gene, such as in assays to identify compounds which modulate expression
30 thereof. In one such assay, the upstream genomic region is ligated to a reporter gene, and incorporated into an expression plasmid. The plasmid is transfected into a cell, and the recombinant cell is exposed to test compound(s). Those compounds which increase or decrease the expression

of the reporter gene are then modulators of the gene/protein, and are considered therapeutic agents of the invention.

The invention also claims recombinant cells engineered to express a polynucleotide or polypeptide as disclosed herein. The gene disclosed herein as being involved in HSN2 in an animal can be used, or a fragment thereof can be used, as a tool to express a protein, where such genes encode a protein, in an appropriate cell *in vitro*, or can be cloned into expression vectors which can be used to produce large enough amounts of protein to use in *in vitro* assays for drug screening. Alternatively, the expression construct may employ the genomic promoter region of HSN2 and link it to a gene, such as a reporter gene, whose expression level is easily measured. Expression systems which may be employed include baculovirus, herpes virus, adenovirus, adeno-associated virus, bacterial systems, and eukaryotic systems such as CHO cells. Naked DNA and DNA-liposome complexes can also be used. The invention thus claims recombinant cell lines containing a heterologous HSN2 gene.

For general molecular biology procedures useful in practicing the present invention, a number of standard references are available that contain procedures well known in the art of molecular biology and genetic engineering and which procedures need not be further described herein. Useful references include Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, N.Y., (1989), Wu et al, *Methods in Gene Biotechnology* (CRC Press, New York, NY, 1997), and *Recombinant Gene Expression Protocols*, in *Methods in Molecular Biology*, Vol. 62, (Tuan, ed., Humana Press, Totowa, NJ, 1997), the disclosures of which are hereby incorporated by reference.

Such recombinant cells may be used in expression assays for analyzing the levels of expression of HSN2 or a suitable reporter gene after contacting said cells with agents that may have analgesic properties. The levels of gene expression can be quantified by Northern blot analysis or RT-PCR, or, alternatively, by measuring the amount of protein produced, by one

of a number of methods known in the art, or by measuring the levels of biological activity of polypeptides encoded thereby or other genes. In this way, the gene expression can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be
5 determined before, and at various points during, treatment of the individual with the agent.

Recombinant cell lines are also preferred for the preparation of purified protein, if a purified protein assay is desired. Those skilled in the art are
10 capable of producing recombinant cell lines and extracting protein fractions containing highly purified proteins. These samples can be used in a variety of binding assays to identify compounds which interact with the proteins. Compounds that interact are therapeutic agents of the invention, or analogs thereof.

15 Target selectivity is an important aspect of the development of therapeutic agents. The present invention specifically contemplates the identification of chemical agents, especially small organic molecules, that inhibit the expression of HSN2, as defined herein, or the activity of the
20 sensorin polypeptide (such as SEQ ID NO: 2 from humans) encoded thereby, with high specificity and that have little or no effect on other genes and/or polypeptides.

Thus, in one such preferred embodiment, the methods disclosed herein
25 for identifying an agent that modulates, preferably inhibits, expression of a gene corresponding to the HSN gene, preferably having the sequence of SEQ ID NO: 1 from humans, or on the activity of a polypeptide encoded thereby, comprises first identifying such agent and then testing such agent for effects on expression or activity of at least one other gene or polypeptide,
30 preferably a gene or polypeptide with important physiological consequences that are preferably not disturbed by therapeutic intervention, and demonstrating little or no effect.

In another aspect, the invention provides a method for computationally identifying a compound of the invention. The method involves (a) determining the active site of a-sensorin protein (i.e. through X-Ray crystallography or other techniques); and (b) through computational modeling, identifying a compound which interacts with the active site, thereby identifying a compound, or its analog, as a compound which is useful for modulating the activity of such a polypeptide. This process is sometimes referred to as *in silico* screening. Sophisticated software for testing the probability of test compounds to interact with the target protein, which can test tens of millions of computer generated compounds, is available to those skilled in the art.

Potential therapeutic compounds identified using the methods of the invention are usually tested in animal model systems to confirm the putative efficacy. Thus, in a further aspect, the present invention relates to a method for identifying an analgesic agent, comprising:

- (a) administering to an animal an agent found to modulate HSN2 gene or protein activity, and
 - b) determining in said animal a decrease in response to a pain stimulus as a result of said administering,
- wherein a determined decrease in response to said pain stimulus indicates analgesic activity, thereby identifying said agent an analgesic agent.

Preferably, the animal is a mammal, such as a human being. In specific embodiments, the pain stimulus is a heat stimulus and reaction or sensitivity to hot and/or cold may be measured. In another embodiment, an electrical stimulus may be used. In all cases, the stimulus may be represented as a sharp or dull sensation. In some cases, the animal may otherwise react normally to such stimulus so that a decrease in normal response due to the test agent is being measured whereas in other cases the animal may initially possess a heightened sensitivity to the stimulus prior to administering the test agent. In all cases, observation of an analgesic effect need not necessarily involve a reduced sensitivity or response to pain but may involve simply a reduced sensation of a particular stimulus. The analgesics identified by the

methods of the invention may induce general analgesia in an animal or may have more localized analgesic or anesthetic effects.

5 In a further aspect, the present invention relates to a method for treating a condition in an animal afflicted with a source of chronic pain comprising administering to said animal an effective amount of an analgesic agent first identified by an assay method of the invention. Preferably, said animal is a human patient, such as a patient afflicted with a chronic ailment, such as a cancerous condition.

10

The screening assays of the invention thus simplify the evaluation, identification and development of therapeutic agents for the treatment of pain, neuropathy and related disorders.

15 The invention also includes antibodies and immuno-reactive substances which target, interact with or bind to sensorin or epitopes thereof. Polypeptides encoded by the polynucleotides disclosed herein can be used as an antigen to raise antibodies, including monoclonal antibodies. Such antibodies will be useful for a wide variety of purposes such as for therapeutic
20 agents, in functional studies, in drug screening assays, in clinical trials and for diagnostics.

For example, in drug screening assays or in clinical trials, the effectiveness of an agent determined by a screening assay as described
25 herein to increase or decrease gene expression, protein levels, or biological activity can be monitored using antibodies specific for sensorin. Alternatively, the effectiveness of an agent determined by a screening assay to modulate expression of HSN2, as well as structurally and functionally related genes, including genes with high homology thereto, and including protein levels, or
30 biological activity can be monitored in clinical trials of subjects exhibiting decreased altered gene expression, protein levels, or biological activity. In such clinical trials, the expression or activity of the genes or polypeptides disclosed herein and, preferably, other genes that have been implicated in, for

example, congenital resistance to pain stimuli, can be used to ascertain the effectiveness of a particular analgesic drug.

In a preferred embodiment, the present invention provides a method for
5 monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, antibody, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) determining that a patient exhibits discomfort due to a disease or disorder that causes some type of painful
10 stimulus; (ii) administering an effective amount of an agent identified using one of the screening assays disclosed herein; (iii) ascertaining a reduction to pain or other stimuli following said administration and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to decrease the expression or
15 activity of gene or encoded polypeptide, i.e., to increase the effectiveness of the agent.

Where the patient is non-human, biopsy samples can be taken to show a decrease in gene expression, such as by measuring levels of protein,
20 mRNA, or genomic DNA post-administration samples and comparing the level of expression or activity of said protein, mRNA, or genomic DNA in the pre-administration sample with that of the corresponding post administration sample or samples, thereby showing the effects of drug administration on one or more of the genes disclosed herein and concomitant reduction in pain
25 response and/or sensitivity.

Candidate modulators may be purified (or substantially purified) molecules or may be one component of a mixture of compounds (e.g., an extract or supernatant obtained from cells). In a mixed compound assay, gene
30 expression is tested against progressively smaller subsets of the candidate compound pool (e.g., produced by standard purification techniques, e.g., HPLC or FPLC) until a single compound or minimal compound mixture is demonstrated to modulate gene or protein activity or expression in a manner having analgesic effects.

Specific compounds which will modulate HSN2 gene expression, gene transcript levels, and protein levels, in a cell include antisense nucleic acids, ribozymes and other nucleic acid compositions which specifically hybridize
5 with HSN2 (including exons or introns of such genes, promoters, 3'-tails, and the like). These specific compounds are compounds of the invention, and are useful for treating the diseases discussed previously. Design and manufacturing of such compounds are well known to those skilled in the art.

10 Specific compounds which modulate the activity of HSN2, more specifically of sensorin, include antibodies (polyclonal or monoclonal) which bind specifically to an epitope of said polypeptide. These antibody compositions are compounds of the invention, and are useful for treating the diseases previously discussed. Design and manufacturing of such
15 compounds are well known to those skilled in the art.

Specific compounds which modulate the activity of HSN2, or its encoded protein, the body include gene therapy vectors comprising all or a part of the HSN2 gene sequence or a mutant HSN2 sequence. As is well
20 known to those skilled in the art, gene therapy allows the delivery of HSN2 in an organism to cells where it is taken up and expressed, thus changing the level or amount of sensorin protein in such cell. These vectors thereby modulate the activity of HSN2 or sensorin in the body and are useful for the therapeutic indications disclosed herein.

25

Specific compounds which modulate the activity of HSN2 or sensorin in the body include small organic molecules. Such compounds may be naturally occurring, or they may be synthetic. Collections and combinatorial libraries of such compounds are widely available from commercial sources. As know to
30 those skilled in the art, a screening assay, such as the assays disclosed in the instant specification, can be easily adapted to identify therapeutic agents which have the desired HSN2 or sensorin modulating ability. Agonists, antagonists, or mimetics found to be effective at reducing response to pain stimuli may be confirmed as useful in animal models (for example, mice,

chimpanzees, etc.). In other embodiments, treatment with a compound of the invention may be combined with other analgesics to achieve a combined, possibly even synergistic, effect.

5 Purified or semi-purified HSN2-encoded protein, or sensorin, and any biochemically modified versions thereof, are themselves therapeutic agents of the invention. Recombinant or non-recombinant forms of these proteins or fragments can be administered to persons in need thereof for the treatment of disorders, such as disorders disclosed herein. Preferably, such agents are
10 administered in a pharmaceutically acceptable carrier. Those skilled in the art are familiar with techniques for generating such agents, and for determining conditions of administration.

Lead Optimization and Analog Development and Selection

15

 In general, novel drugs having analgesic or anesthetic properties, and which modulate sensorin activity, are identified from libraries, possibly large libraries, of both natural product or synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. Those skilled in the
20 field or drug discovery and development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the exemplary methods described herein. Examples of such extracts or compounds include, but are
25 not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (e.g., semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-,
30 lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex,

UK), Xenova (Slough, UK), Harbor Branch Oceanographics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, e.g., by standard extraction and fractionation
5 methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

De-replication(e.g., taxonomic dereplication, biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of
10 replicates or repeats of materials already known for their analgesic and/or anesthetic activities may be employed whenever possible.

When a crude extract is found to have analgesic and/or anesthetic activities, or both, further fractionation of the positive lead extract is possible
15 to isolate chemical constituent responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having such analgesic and/or anesthetic activities. The same assays described herein for the detection of activities in mixtures of compounds can
20 be used to purify the active component and to test derivatives thereof. Methods of fractionation and purification of such heterogeneous extracts are known in the art. If desired, compounds shown to be useful agents for the treatment of pathogenicity are chemically modified according to methods known in the art. Compounds identified as being of therapeutic value are
25 subsequently analyzed using any standard animal model of diabetes or obesity known in the art.

In general, these screening methods provide a ready means for selecting either natural product extracts or synthetic compounds of interest
30 from a large population (i.e. a chemical library, for example, one produced by combinatorial means) which are further evaluated and condensed to a few active core structures. Multiple analogs of such core structures may be developed and tested to identify those preferred analogs which have improved characteristics as therapeutic agents.

Improved analogs may also include compounds with improved stability, biodistribution, pharmacokinetics or other desirable features for therapeutic agents which are not directly related to modulation of the therapeutic target.

5 In a preferred embodiment, the improved analog of the invention is effectively delivered, either by physiological means or assisted means, to cells of the body expressing the sensorin protein. Analog compounds are systematically screened to evaluate whether they modulate the identified biological activity and those that effectively do so are then therapeutic agents, or further
10 analogs thereof, according to the invention.

Therapeutic Agents and Uses Thereof

The agents contemplated by the present invention are highly selective
15 for the HSN2-encoded protein (sensorin) and administration of such an agent to a human or other animal in need thereof provides a treatment for any of the disorders exemplified by, or later found to be related to, HSAN II (and involving possible mutations of HSN2). For example, an inhibitor mimics the effects of one or more of the mutated forms of the gene as disclosed herein.
20 Alternately, an agonist would reverse a disorder made evident in HSAN II. HSN2 and/or sensorin may also function in the development, survival or maintenance of peripheral sensory neurons, or supporting cells and thus may be useful in identifying agents that act directly to further the life and activities of such cells. Genetic or pharmacological down-regulation of this protein may
25 lead to novel analgesics used to treat pain or other neurological symptoms, whereas up-regulation may lead to novel neuronal survival agents which may be useful in treating a number of neurodegenerative disorders of both the CNS and PNS. As such, the agent identified according to this invention is a treatment for pain, neuropathy or a related disorder.

30

Disorders which may be related to aberrant HSN2 or sensorin activity include pain, for example inflammatory pain, a neuralgia, a nerve entrapment syndrome, and pain associated with a musculoskeletal disorder;

neuropathies, for example peripheral, diabetic, autonomic, sensory or motor neuropathy; control of sweating; autonomic dysfunction such as control of blood-pressure, swallowing, posture, hypertension, hypotension, GI – esophageal dysmotility, impotence, apnea, absence of lacrimation; Riley-Day
5 Syndrome, temperature control, hypertonia; delayed gastric emptying, diarrhea, vomiting, neurological disorders such as seizure, migraine, tremor, anxiety, depression, hyperphagia and associated obesity, bipolar disorder, Parkinson's, Huntington's, spinal cord injury, multiple sclerosis, traumatic brain injury, stroke, neurodegenerative disorders such as Alzheimer's Disease
10 and ALS, and the like.

A variety of diseases are also treatable based on sensorin's possible role as a novel neurotrophic growth factor in the nervous system. Sensorin's amino acid sequence is about twice the length of most other known
15 neurotrophic factors, and may represent a proform that is proteolytically cleaved at positions other than at the signal peptide cleavage site. Like other neurotrophic factors, sensorin may be expressed at low levels.

Neurotrophic factors are essential for the development of the nervous
20 system. They exert a diverse set of influences including neuronal cell survival, axonal and dendritic growth and guidance, synaptic structure and connections, neurotransmitter release, long-term potentiation and synaptic plasticity. Alterations in neurotrophic factor levels can influence cell survival and apoptosis during development, myelination, regeneration, pain,
25 aggression, depression, substance abuse, anxiety, hyperphagia, memory acquisition and retention and loss of the peripheral nervous system. It is thus suggested that administration of sensorin protein, or fragments thereof, or modulators of sensorin or HSN2 activity are useful for the treatment of these diseases.

30

Those skilled in the art are familiar with the necessary steps for pre-clinical and human clinical trials which are used to establish efficacy and safety of the new chemical entities and compounds first identified by the invention for use in treating the diseases mentioned herein.

Compounds first identified as useful in reducing sensitivity to pain stimuli using one or more of the assays of the invention may be administered with a pharmaceutically-acceptable diluent, carrier, or excipient, in unit dosage form. Conventional pharmaceutical practice may be employed to provide suitable formulations or compositions to administer such compositions to patients. Although oral administration is preferred, any appropriate route of administration may be employed, for example, intravenous, parenteral, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intrathecal, epidural, intracisternal, intraperitoneal, intranasal, or aerosol administration. Therapeutic formulations may be in the form of liquid solutions or suspension; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.

Methods well known in the art for making formulations are found in, for example, Remington: The Science and Practice of Pharmacy, (19th ed.) ed. A.R. Gennaro AR., 1995, Mack Publishing Company, Easton, PA. Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated naphthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for agonists of the invention include ethylenevinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, or example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.

30

Combination therapies are also contemplated by the inventors. An analgesic agent identified by one of the screening methods disclosed herein may be administered along with another agent intended to treat a coincident

conditions, such as where analgesic and antitumor agents are given together or contemporaneously.

5 The present invention also relates to a process that comprises a method for producing a product, such as by generating test data to facilitate identification of such product, comprising identifying an agent according to one of the disclosed processes for identifying such an agent (i.e., the therapeutic agents identified according to the assay procedures disclosed herein) wherein said product is the data collected with respect to said agent
10 as a result of said identification process, or assay, and wherein said data is sufficient to convey the chemical character and/or structure and/or properties of said agent. For example, the present invention specifically contemplates a situation whereby a user of an assay of the invention may use the assay to screen for compounds having the desired enzyme modulating activity and,
15 having identified the compound, then conveys that information (i.e., information as to structure, dosage, etc) to another user who then utilizes the information to reproduce the agent and administer it for therapeutic or research purposes according to the invention. For example, the user of the assay (user 1) may screen a number of test compounds without knowing the
20 structure or identity of the compounds (such as where a number of code numbers are used the first user is simply given samples labeled with said code numbers) and, after performing the screening process, using one or more assay processes of the present invention, then imparts to a second user (user 2), verbally or in writing or some equivalent fashion, sufficient
25 information to identify the compounds having a particular modulating activity (for example, the code number with the corresponding results). This transmission of information from user 1 to user 2 is specifically contemplated by the present invention.

30 In accordance with the foregoing, the present invention encompasses a method for producing test data with respect to the gene modulating activity of a compound comprising:

(a) contacting a test compound with a HSN2 gene under conditions where said gene is being expressed;

b) determining a change in the expression of said gene as a result of said contacting, and

(c) producing test data with respect to the gene modulating activity of said test compound based on a change in the expression of the determined
5 gene as a result of said contacting.

Diagnostics and Pharmacogenomics

In a further embodiment, the invention relates to diagnostic and
10 pharmacogenomic compounds, kits and methods. This aspect relates to analysis HSN2 for the diagnosis of insensitivity or indifference to pain, other pain disorder, or in the selection of a therapeutic agent for a patient (i.e. pharmacogenomics).

15 For example, nucleic acid analysis can be used to identify the HSN2 mutations disclosed herein, thus confirming the diagnosis of HSN II. Many nucleic acid diagnostic techniques are well known to those skilled in the art. Such techniques include DNA sequencing, hybridization probing, single stranded conformational analysis, PCR based techniques such as mismatch
20 amplification, and myriad other well known methods. All such analysis can be performed on a small sample of blood, saliva, urine or other tissue provided by the patient.

Alternatively, protein based analyses such as antibody based assays
25 (Elisa, Radioimmunoassay and the like) can be employed to identify the expression, amount or presence or absence of a mutant protein (such as sensorin encoded by a mutant HSN2), such as those disclosed herein.

Gene expression, both comparable and absolute, as well as biological
30 activity, and mutational analysis can each serve as a diagnostic tool for pain or neuropathic disorders; thus determination of the amount of HSN2 mRNA can be used to diagnose the presence or absence of a mutation correlated with such pain or neuropathic disorder.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons (Eichelbaum, M., Clin. Exp. Pharmacol. Physiol., 23:983-985, 1996; Linder, M. W., Clin. Chem., 43:254-266, 1997). In general, two
5 types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). Altered drug action may occur in a patient having a polymorphism (e.g., an single nucleotide
10 polymorphism or SNP) in promoter, intronic, or exonic sequences of HSN2. Thus by determining the presence and prevalence of polymorphisms in HSN2 in an individual, one may predict a patient's response to a particular therapeutic agent.

15 This pharmacogenomic analysis can lead to the tailoring of drug treatments according to patient genotype, including prediction of side effects upon administration of therapeutic agents, particularly therapeutic agents for treating disorders disclosed in this specification. Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic
20 treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual is examined to determine the ability of the individual to respond to a particular agent).

25 Diagnostics employing a gene or protein corresponding to HSN2 can also be useful in selecting patients for clinical trials of a potential therapeutic agent. Patients can be stratified according to the DNA or protein sequence of HSN2 and their response to drug treatment can be evaluated. Such stratification can greatly reduce the number of patients required to establish efficacy for a potential therapeutic agent.

30

Example 1

Identification of the genetic mutation responsible for Hereditary Sensory and Autonomic Neuropathy – Type II (HSAN II).

5

We performed a genomewide screen in two geographically isolated families, who are likely related, to map the HSAN II locus. We expanded a consanguineous multigenerational family with eight affecteds (HSAN4) from a sibship reported previously and collected a smaller family with two affecteds (HSAN3). Most of the patients live within a 100-mile radius in a geographically isolated region. In both families the mode of inheritance was autosomal recessive and affected members presented with severe and early-onset HSAN II. Sensory loss was predominantly distal, but the progression of the disorder varied, involving the trunk in some patients. Because of the common geographic origin of our subjects and the relationships established through family history, it became clear that the different degrees of sensory loss were part of a continuum of a single disease process, and both families were expected to segregate the identical mutation. Pedigrees are shown in Fig 1. Additionally, unrelated affected samples from a different population were available, and these consisted of two affected sisters (HSAN5-301 and HSAN5-302) and another affected individual with no close family ties to the affected sisters (HSAN6-301).

25 Fifteen family members from HSAN4 and seven family members from HSAN3 were genotyped at 763 autosomal markers and at 48 X markers. A genome-wide two-point analysis identified five regions with cumulative two-point LOD scores >2.0 . Only one region contained a marker that was homozygous in the three affected siblings in HSAN4 (patients 70-72) and the two affected siblings in HSAN3 (patients 31-42). This marker, D12S352, with a positive cumulative LOD score of 2.64 at zero recombination, suggested that HSAN type II likely maps to chromosome 12p13. In addition, this was the only genomic segment where a three-marker haplotype was shared between all affecteds. Analysis of additional microsatellite markers demonstrated that

all affected individuals in both families were homozygous for each marker in a 1.2 Mb common haplotype. Marker order was consistent between the Genethon and Marshfield genetic maps and the Build 30 human genome sequence assembly, suggesting that there were no gross inconsistencies or recombination hotspots. 22 markers gave significant positive cumulative LOD scores, the highest one being 8.4 for marker CA1AC005343 at zero recombination.

The core region of shared marker alleles (see Table 1) places the HSAN II locus within a region on chromosome 12p13 between markers CA1AC0021054 (forward primer: 5'-ACCATCACCTAAGGAGACAGACC (SEQ ID NO: 30) and reverse primer: 5'-TGCAACAAATGTACCACTCTGG (SEQ ID NO: 31)) and D12S1642 (forward primer: 5'-AGCTCCTAAATCCCCG (SEQ ID NO: 32), and reverse primer: 5'-GCCATGTCTATAAATACCCTG (SEQ ID NO: 33)) based on recombination breakpoints in individuals HSAN4-171 and HSAN4-124.

We also screened the additional affecteds from a different population. Individual HSAN6-301 was homozygous for all markers in the region defined in the HSAN4 family. The two affected sisters, HSAN5-301 and HSAN5-302, were not homozygous in the 12p13 region, and instead shared alleles on one chromosome with individual HSAN6-301. Alleles on the other chromosome are shared between the two sisters, suggesting that they are compound heterozygotes and that there are two founder mutations in this population. A recombination in one of the HSAN5 sisters positions the linked region above marker CA1AC005183x2.

The genomic sequence for the region defined in family HSAN4 is well defined, with a BAC contig consisting of 12 completely finished and one incomplete BAC sequences. The syntenic region in mouse is ~900 kb in size and located on chr 6. The gene order is conserved between human and mouse.

Table 1. Definition of the HSN II minimal candidate region

Marker	Genethon (cM)	Marshfield (cM)	June2002 (kb)	HSAN3-31-a	HSAN3-31-b	HSAN3-42-a	HSAN3-42-b	HSAN4-57-a	HSAN4-57-b	HSAN4-70-a	HSAN4-70-b	HSAN4-171-a	HSAN4-171-b	HSAN4-109	HSAN4-122/124	HSAN5-301-a	HSAN5-301-b	HSAN5-302-a	HSAN5-302-b	HSAN6-301-a	HSAN6-301-b
GGAT2AC007406				325	297	297	297	317	325	297	297	321	317	297		234	246	234	246	236	236
D12S352	0	0	181	154	154	154	154	154	160	154	154	166	154								
GAAA2AC021054			573	369	369	369	369	369	369	369	369	377	369	369							
GAAA1AC021054			632	216	216	216	216	216	216	216	216	211	216	216	216						
CA1AC021054			635	246	246	246	246	246	246	246	246	234	246	246							
CA2AC021054			639	237	237	237	237	237	237	237	237	237	237	237	237						
GAAA3AC021054			684	437	437	437	437	437	433	433	433	437	437	433	433						
D12S341	0.6	0.62	702	122	122	122	122	122	122	122	122	122	122	122	122	123	119	123	119	123	123
CA1AC004765			738	174	174	174	174	174	174	174	174	174	174	174		191	189	191	189	191	191
CA2AC004765			751	281	281	281	281	281	281	281	281	281	281	281		281	281	281	281	281	281
D12S94	1.1	1.17	763	201	201	201	201	201	201	201	201	201	201	201	201	201	201	201	201	201	201
D12S91	1.1	1.71	800	181	181	181	181	181	181	181	181	181	181	181	181	179	184	179	184	179	179
CA1AC004803			859	249	249	249	249	249	249	249	249	249	249	249	249	255	251	255	251	255	255
A1AC004803			949	199	199	199	199	199	199	199	199	199	199	199		206	204	206	204	205	206
D12S389	0	1.71	984	164	164	164	164	164	164	164	164	164	164	164		156	161	156	161	156	156
D12S1285	0	1.71	1403	300	300	300	300	300	300	300	300	300	300	300		308	308	308	308	308	308
D12S1608	2.7	2.79	1628	250	250	250	250	250	250	250	250	250	250	250		246	248	246	248	246	246
CA1AC005182			1630	217	217	217	217	217	217	217	217	217	217	217	217	207	207	207	207	207	207
D12S1656	2.7	2.79	1679	271	271	271	271	271	271	271	271	271	271	271		269	269	269	269	269	269
CA1AC005183x2			1692	290	290	290	290	290	290	290	290	290	290	290		288	288	288	290	288	288
CA3AC005343			1745	305	305	305	305	305	305	305	305	305	305	305		297	297	297	303	297	297
CA2AC005343			1783	245	245	245	245	245	245	245	245	245	245	245		253	253	253	253	253	253
CA1AC005343			1785	194	194	194	194	194	194	194	194	194	194	194	194	185	183	185	183	185	185
CA1AC0090840			1878	190	190	190	190	190	190	190	190	190	190	190	190	188	188	188	188	188	188
D12S1642	3.3	3.33	1906	159	159	159	159	159	159	159	159	159	159	159	161	188	188	188	188	188	188
CA2AC005342			1975	381	381	381	381	381	381	381	381	381	381	381	371	188	198	188	190	188	190
CA1AC005342			2037	197	197	197	197	197	197	197	197	197	197	197							
D12S100	3.3	3.33	2048	153	153	153	153	153	153	153	153	153	153	153	151						
D12S1689	3.3	3.33	2207	103	103	103	103	103	103	103	103	103	103	103	109						
D12S1694	3.3	3.33	2258	363	363	363	363	363	363	363	363	363	363	363							
D12S1615	4.6	4.64	2642	nd	nd	nd	nd	nd	215	215	215	215	215	215	213						
D12S1626			3034	188	188	188	188	188	188	188	188	188	188	188							

candidate region defined in pedigrees HSN3 and HSN4

candidate region defined in pedigrees HSN5 and HSN6

Candidate genes: A list of the genes mapping to this candidate interval was compiled based on the Apr2002 freeze of the human genome assembly available from UCSC. A total of 13 known or predicted transcripts are located between markers D12S352 and D12S1615. These included NINJ2 (NM_016533), FLJ31553 (AK056115), PRKWNK1 (NM_018979), RAD52 (NM_002879), ELKS (NM_015064), PPA (BM714852), WNT5B (NM_030775), FLJ21432 (BM453101), CACNA2D4 (AL137658), SLIT2L (BI754042), FLJ25180 (AK057909), FLJ31638 (AK056200) and CACNA1C (NM_000719). The transcripts were categorized according to their evidence for being real genes, according to the following scheme: 1 - confirmed mRNA; 2 - mRNA supported by ESTs; 3 - spliced ESTs with other evidence to support; 4 - single spliced EST; 5 - unspliced EST; 6 - pseudogene not expressed. Many unspliced ESTs were found, and no serious effort was taken to label these as transcripts, given the known contamination of genomic DNA or unspliced intronic contamination within the cDNA libraries used to generate dbEST. Also, gene predictions were generally ignored unless there was independent evidence from mRNA sequence, EST, or non-human gene in which case the mRNA was used as transcript information.

Sequencing of the coding regions of these initial 13 candidate genes did not identify any relevant mutations. A more exhaustive search for candidate genes was therefore conducted based on conserved homologies between human and mouse genomic assemblies. The human genomic sequence representing the entire HSN II candidate interval (1.2 Mb) was downloaded from the Human Genome Browser website (genome.ucsc.edu/) with case toggled to highlight the mouse translated BLAT track (which represents regions showing significant protein homology between human and mouse). A similar human sequence showing the translated Tetraodon BLAT track was also downloaded. These sequences were assembled into a contig, to which all exons of previously identified candidate genes were added. Then the contig was scanned for novel conserved fragments, roughly >80% conserved over >100 bp. This identified 64 novel fragments, which were then tested for functional homology using: 1) BLASTn against the nr database, 2) BLASTn against the dbEST, and 3) BLASTx against the nr database available

from the NCBI website (at www.ncbi.nlm.nih.gov/BLAST/). The HSN2 gene contained the only novel well-conserved ORF identified in this screen.

5 Assembly of HSN2 cDNA contig and transcripts: A total of 21 human EST sequences (from 8 different cDNA libraries), including BG428619, BF522762, and N98701, and a genomic sequence from BAC AC004765, were assembled into a contig. A putative cDNA sequence was derived from the genomic sequence, with ends based on human ESTs (SEQ ID NO. 1). The start ATG is immediately preceded by a stop TGA codon, and there is no
10 further indication of a conserved splicing site upstream of the ORF, suggesting that this gene consists of a single unspliced exon. The identified open reading frame (ORF) encoding a peptide of 434 aa (SEQ ID NO. 2) was used to identify the mouse, rat, and fish orthologs from their respective genome assemblies. A consensus genomic sequence was downloaded from
15 the UCSC website (SEQ ID NO. 12). The HSN2 gene maps to intron 8 of the PRKWINK1 gene, and consists of a single exon that is transcribed from the same strand as PRKWINK1.

To screen the ORF for nucleic acid mutations, three separate
20 amplicons were designed (XH18x01a, XH18x01b, and XH18x01c). PCR primers (XGR0720 to XGR0725) were designed using PrimerSelect (DNASTAR, Madison WI) and purchased from BioCorp (Montreal). The primer sequences, and amplicon sizes are given in Table 2. The TD1 protocol for PCR amplification consisted of the following: initial denaturation for 5 min. at
25 94C, followed by 17 cycles of denaturation at 95°C for 30 sec, annealing at a temperature starting at 70°C and ending at 54°C (-1°C per cycle), and 45 sec elongation at 72C, followed by 25 cycles of denaturation at 95°C for 30 sec., annealing at 54°C for 30 sec. and elongation at 72°C for 45 sec., followed by a single cycle at 72°C for 5 min.

Table 2.

Amplicon	Forward primer	Reverse primer	Size (bp)	Protocol
XH18x01a	XGR0720: (SEQ ID NO: 24) TTCCAGAAGCATTTGTTATTTATTT	XGR0721: (SEQ ID NO: 25) CCCCCTTGTAAGGCTTCT	677	TD1
XH18x01b	XGR0722: (SEQ ID NO: 26) CACCAGAGGCCGTAGTTATGTTG	XGR0723: (SEQ ID NO: 27) TTGAGGAGGCAGTTCTTCTTGATT	638	TD1
XH18x01c	XGR0724: (SEQ ID NO: 28) GCGCCTGCTGTGTAACTCATAA	XGR0725: (SEQ ID NO: 29) CCAAAGATGGGGAACTCTACTGA	669	TD1

Samples were amplified using Taq polymerase (Qiagen) on a PE 9700 thermocycler. All amplified fragments were tested by agarose gel electrophoresis. Samples tested were HSN4-70 (affected homozygous), HSN4-371 (normal control from same pedigree) and HSN5-302 (affected). Amplified products and sequencing primers were sent to the Montreal Genome Centre sequencing facility for sequence determination. Traces were then aligned using a genomic sequence contig constructed in SeqManII (DNASTAR), and variations identified and annotated to an Excel spreadsheet. Publicly available predicted SNPs (from NCBI dbSNP) were identified, or those found in the cDNA contig were noted. Sequence variants were documented by printing the screen image in MS PhotoEditor, selecting the relevant sequence trace and pasting into MSWord to make a Figure.

15

Identification of mutations in HSN2: Three separate mutations were identified that were predicted to gravely affect the ORF of the HSN2 gene (see Table 3, SEQ ID NO: 3 to 5). A deletion of an A at position 594 (relative to the start ATG), found homozygous in the HSN4-70 sample, is predicted to cause a frameshift in codon 198 leading to premature truncation of a 206 aa peptide. Two different mutations were found in the HSN5-302 sample. The first is a 1 bp insertion of an A between bases 918-919, causing a frameshift in codon 307, and leading to premature truncation of a 318 aa peptide. The second is a C>T change at position 943 that changes codon 315 (CAG, encoding Glutamine) to a TAG stop codon, therefore prematurely truncating the protein at 314 amino acids. Each of these mutations was not found in the control sample, dbSNP database, nor available genomic and EST sequences.

25

Table 3. Sequence ID Numbers

SEQ ID NO.	Description
1	Wild-Type Human HSN2 putative cDNA (ends based on human ESTs)
2	Wild-Type Human sensorin
3	Mutation 1 - Amino acid sequence of truncated human sensorin in HSN4-70
4	Mutation 2 - Amino acid sequence of truncated human sensorin in HSN4-302
5	Mutation 3 - Amino acid sequence of truncated human sensorin in HSN4-302
6	Genomic promoter sequence for human HSN2
7	Wild-Type Mouse HSN2 cDNA predicted (5' and 3' end not determined)
8	Wild-Type Mouse sensorin amino acid sequence. 433 amino acids
9	Wild-type Rat HSN2 nucleic acid sequence. (genomic sequence defined by 5' and 3' ends in EST BF522762/AI578184, contains polyA tail)
10	Wild-type rat sensorin amino acid sequence. 434 amino acids
11	Wild-type Fugu sensorin amino acid sequence
12	Genomic structure and sequence of the human HSN2
13	Partial wild-type amino acid sequence of Tetraodon sensorin
14	Human cDNA; mutation 1, c.594delA
15	Human cDNA; mutation 2, c.918insA
16	Human cDNA; mutation 3, c.943C>T
17	Pig HSN2 Nucleotide Sequence
18	Pig HSN2 Amino Acid Sequence
19	Human HSN2 translated from 2 nd ATG
20	Human HSN2 translated from 3 rd ATG
21	Mouse HSN2 translated from 2 nd ATG
22	Mouse HSN2 translated from 3 rd ATG
23	Rat HSN2 translated from 2 nd ATG

5 SEQ ID NO: 17 is the complete sequence of pig (*Sus scrofa*) cDNA clone, which was determined from clone MI-P-AY1-nrg-c-02-0-UI generated by Pig Genome Project and partially sequenced with Genbank accession number BI399422. SEQ ID NO: 18 is the sequence for the pig HSN polypeptide.

10

Analysis of SEQ ID NO: 19 to 23 shows that the 3rd ATG is closest to the optimal Kozak consensus C-C-[AG]-C-C-A-T-G-G context for initiation of translation, but is not conserved in rat Hsn2 genomic sequence.

WHAT IS CLAIMED IS:

1. A method for identifying an agent that modulates the activity of the HSN2 gene, comprising:

5 a) contacting a test compound with a cell that expresses an HSN2 gene;

 b) determining a change in the expression of said gene as a result of said contacting,

 wherein said determined change in expression of the gene indicates modulation, thereby identifying the test compound as an agent that modulates the activity of an HSN2 gene.

10

2. The method of claim 1 wherein said HSN2 gene is a mammalian HSN2I gene.

15

3. The method of claim 2 wherein the mammal is mouse, rat or human.

4. The method of claim 1 wherein said modulation is a decrease in expression.

20

5. The method of claim 1 wherein said cell is a mammalian cell.

6. The method of claim 5 wherein said cell is a recombinant cell.

25 7. The method of claim 5 wherein said cell is a cell of the nervous system.

8. The method of claim 1 wherein said gene comprises a polynucleotide corresponding to a polynucleotide having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 7 and 9.

30

9. A method for identifying an agent that modulates HSN2 gene, comprising:

(a) contacting a test compound with a genetic construct comprising a reporter gene operably linked to an HSN2 promoter under conditions where the reporter gene is expressed;

5 (b) determining a change in expression of the reporter gene as a result of said contacting,

wherein a determined change in expression indicates modulation, thereby identifying the test compound as an agent that modulates HSN2 gene.

10 10. The method of claim 9 wherein said modulation is a decrease in expression of said reporter gene.

11. The method of claim 9 wherein said HSN2 promoter is a mammalian HSN2 promoter.

15 12. The method of claim 11 wherein the mammal is mouse, rat or human.

13. The method of claim 9 wherein said promoter is the promoter of
20 SEQ ID NO: 6.

14. The method of claim 9 wherein said genetic construct is present in a cell.

25 15. The method of claim 9 wherein said cell is a mammalian cell.

16. The method of claim 9 wherein said cell is a recombinant cell.

17. The method of claim 15 wherein said cell is a nervous system cell.

30 18. A method for identifying an agent that modulates the activity of an HSN2-encoded protein, comprising:

a) contacting a test compound with an HSN2-encoded polypeptide under conditions where said polypeptide is active; and

b) determining a change in the activity of said polypeptide as a result of said contacting;

wherein said determined change in activity indicates modulation, thereby identifying the test compound as an agent that modulates the activity of HSN2-encoded protein.

19. The method of claim 18 wherein said HSN2-encoded protein is a mammalian HSN2-encoded polypeptide.

20. The method of claim 19 wherein the mammal is mouse, rat or human.

21. The method of claim 18 wherein said modulation is a decrease in activity.

22. The method of claim 18 wherein said protein is present in a mammalian cell.

23. The method of claim 22 wherein said cell has been engineered to contain said protein.

24. The method of claim 23 wherein said cell was engineered by genetic engineering.

25. The method of claim 23 wherein said cell does not contain said protein absent said engineering.

26. The method of claim 22 wherein said cell is a cell of the nervous system.

27. The method of claim 22 wherein said protein is encoded by a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1, 7, and 9.

28. The method of claim 22 wherein said protein comprises an amino acid sequence of SEQ ID NO: 2, 8, 10 or 11.

29. A method of treating a pain-related disorder comprising
5 administering to an animal in need thereof a therapeutically effective amount of an HSN2 modulator.

30. The method of claim 29 wherein said HSN2 modulator exhibits modulating activity in an assay method of one of claims 1 to 28.
10

31. The method of claim 29 wherein said agent was first identified as an HSN2 modulator using an assay method of one of claims 1 to 28.

32. A method for treating an HSN2-related disorder comprising
15 administering to a person in need thereof an effective amount of a selective sensorin agonist or antagonist, or a pharmaceutically acceptable salt thereof, or a pharmaceutical composition containing said agonist or antagonist.

33. The method according to claim 32 wherein said disorder is pain,
20 neuropathy or a related disorder.

34. The method of claim 33, wherein said pain disorder is selected from the group consisting of inflammatory pain, a neuralgia, a nerve entrapment syndrome, and pain associated with a musculoskeletal disorder.
25

35. A method of diagnosing the presence of, or risk of developing, an HSN2-related disorder comprising determining the presence of a mutation in the nucleic acid sequence of the HSN2 gene of said individual or in the amino acid sequence of an HSN2-encoded protein of said individual.
30

36. A method for identifying an agent that modulates HSN2 gene activity or sensorin activity, comprising: (a) contacting a test compound with a cell expressing HSN2; and (b) determining a change in expression of a HSN2 nucleic acid or the activity of sensorin, wherein said change indicates

modulation thereby identifying the test compound as an agent that modulates HSN2 gene activity or sensorin activity.

37. The method of claim 36, wherein the test compound is a small
5 molecule.

38. The method of claim 36, wherein the test compound is an anti-sensorin antibody.

10 39. The method of claim 36, wherein the test compound is an antisense HSN2 nucleic acid molecule.

40. The method of claim 36, wherein the test compound is a HSN2 ribozyme.
15

41. An isolated polynucleotide comprising a nucleotide sequence with at least 60% identity to a sequence selected from the group consisting of SEQ ID NO: 1, 6, 7, 9 and 12 wherein said isolated polynucleotide encodes a polypeptide that binds to an antibody specific for a polypeptide having the
20 amino acid sequence of SEQ ID NO: 2.

42. The isolated polynucleotide of claim 41 wherein said identity is at least 70%.

25 43. The isolated polynucleotide of claim 41 wherein said identity is at least 78%.

44. The isolated polynucleotide of claim 41 wherein said identity is at least 90%.
30

45. The isolated polynucleotide of claim 41 wherein said identity is at least 95%.

46. The isolated polynucleotide of claim 41 wherein said identity is at least 98%.

47. The isolated polynucleotide of claim 41 wherein said polynucleotide
5 has a sequence selected from the group consisting of SEQ ID NO: 1, 6, 7, 9 and 12.

48. An isolated polypeptide comprising an amino acid sequence with at least 90% identity to an amino acid sequence selected from the group
10 consisting of SEQ ID NO: 2, 8, 10, 11 and 13 wherein said polypeptide binds to an antibody specific for a polypeptide having the amino acid sequence of SEQ ID NO: 2..

49. The isolated polypeptide of claim 48 wherein said percent identity is
15 at least 95%.

50. The isolated polypeptide of claim 48 wherein said percent identity is at least 98%.

20 51. An isolated polypeptide comprising a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, 3, 4, 5, 8, 10, 11 and 13.

52. A method for identifying an analgesic agent, comprising:
25 a) administering to an animal an agent found to modulate HSN2 gene or HSN2-encoded polypeptide activity, and
b) determining in said animal a decrease in response to a pain stimulus as a result of said administering,
wherein a decrease in response to said pain stimulus indicates
30 analgesic activity, thereby identifying said agent as an analgesic agent.

53. The method of claim 52 wherein said animal is a mammal.

54. The method of claim 53 wherein said mammal is a mouse, a rat or a human being.

55. A method for producing test data with respect to the gene
5 modulating activity of a compound comprising:

(a) contacting a test compound with a HSN2 gene under conditions where said gene is being expressed;

b) determining a change in the expression of said gene as a result of said contacting, and

10 (c) producing test data with respect to the gene modulating activity of said test compound based on a change in the expression of the determined gene as a result of said contacting.

Figure 1A

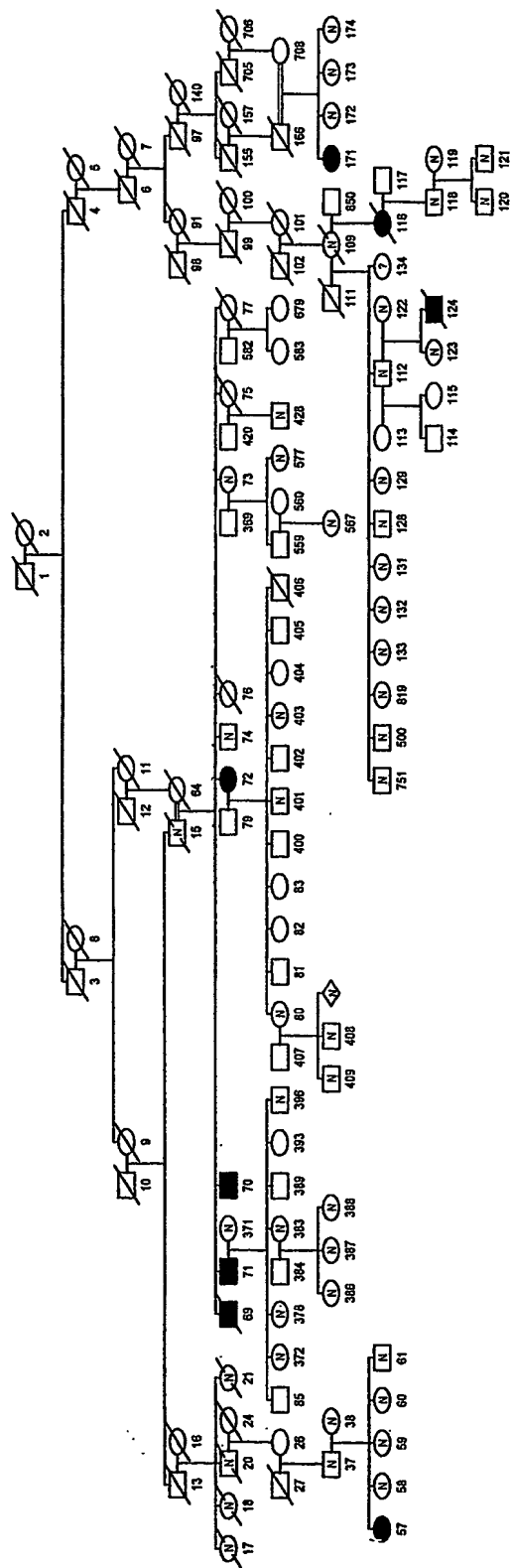


Figure 1B

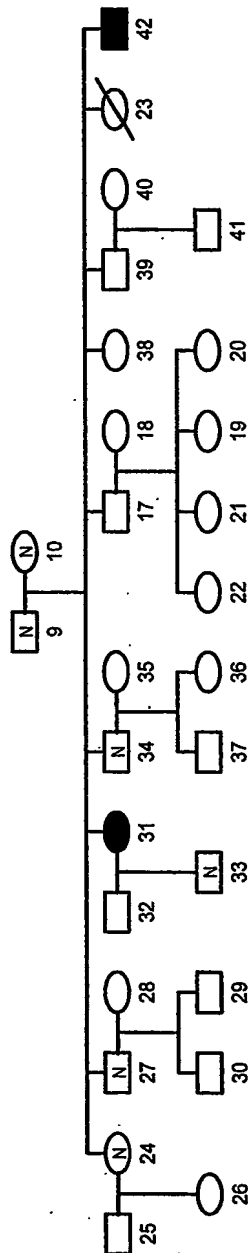
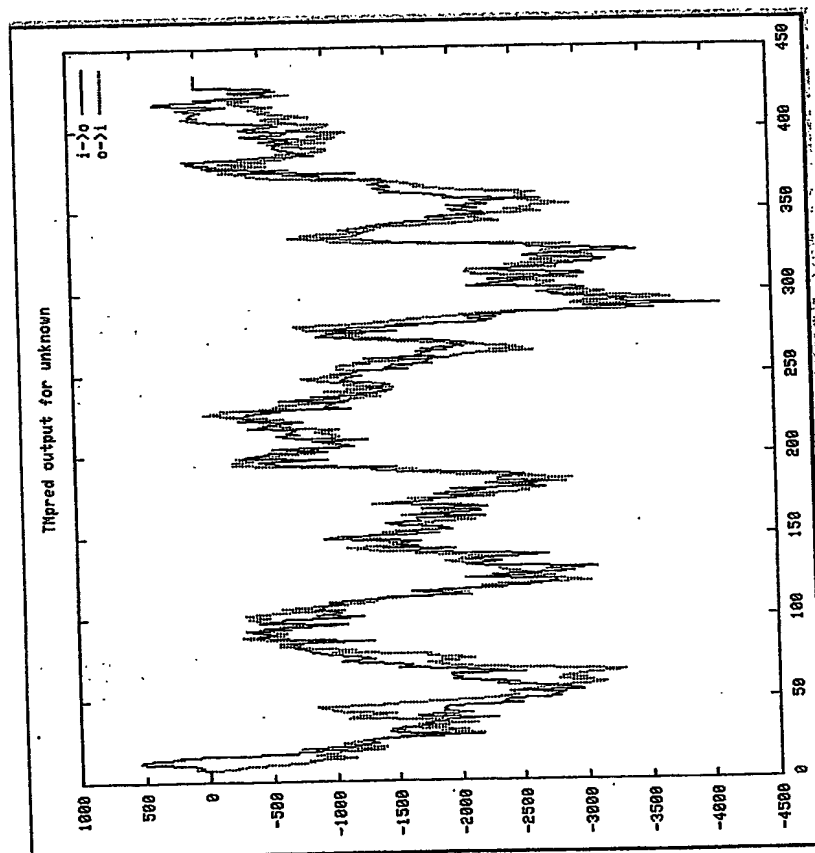


Fig. 2. Domain prediction of the sensorin protein.
Predicted TM topology using TMPred.



Human 1 --MYELVLFMLIQPSMAHP-----CGGTP--TYPE--SOIEFFETIHERPVSFSPPTCPKVAISQRRKSTSFLEAQTTHFQPLL
 Mouse 1 --MYELVLFMLIQPSMAHP-----CGGTP--TYPE--SOIEFFETIHERPVSFSPPTCPKVAISQRRKSTSFLEAQTTHFQPLL
 Rat 1 --MYELVLFMLIQPSMAHP-----CGGTP--TYPE--SOIEFFETIHERPVSFSPPTCPKVAISQRRKSTSFLEAQTTHFQPLL
 Tetra 1 SECHSEHCISLSQHSMPHYSGGAGGAGCNVPLASLPDPPTVLFPSIPERPIFSFPPTMPKAYNNQRRKSTSI LEAHTRHFPAY
 Fugu 1 FECHS* IHCISLSQHSMPHYSGGAGGAGCNVPLASLPDPPTVLFPSIPERPIFSFPPTMPKAYNNQRRKSTSI LEAHTRHFPAY

Human 77 RTVSQSLFPGGSPNTWTEAVVM-----IGTASRVGTGESCEIQVHPMFEPQVSDYRPGLVLPETARYFIPOEAVYVAGVHYQARVAEQ
 Mouse 77 RTVSQSLFPGGSPNTWTEAVVM-----IGTANRVNRELCEMQQVPEFTQIKSDYRPGLVLAEEARYFIPOETVYLAGVHYQAVAGQ
 Rat 77 RTVSQSLFPGGSPNTWTEAVVM-----IGTANRVNRELCEMQQVPEFTQIKSDYRPGLVLAEEARYFILLRETVYLGGVHYHPHAAGQ
 Tetra 91 PRYSSNLHFEFCMEAVEPPLFMVNPFGPSPAQRIG-----VGEPLHPCQSDPSLYGYKDMFAEHGEAVRRLS-----QALLDH
 Fugu 90 TRYSSNLHFEFCMETVEPPLFMVNPFGPSPAQRIG-----VGEPLHPCQSDPSLYGYKDMFAEHGEAMRRSL--NQALLDH

Human 164 YEGIPYNS-SVLSSPMKQIP-EQKPVQCGPTS-----SSVEFEFGQAFLVGHIL--QNIRLDSGLGFCGSPPLSSISAFISTDATRLKFHPV
 Mouse 164 YEGIPYNS-EVLSSPMKQIS-EQKPVPCGPAS-----SSVEFEFGQAFLVGHIL--QNIRLDSGPSFASPLSSISAFNSTDATHLKFHPV
 Rat 164 YEGIPYNS-EVLSSPMKQIT-EQKPVPCGPAS-----SSVEFEFGQAFLVGHIL--QNIRLDSGPSFASPLSSISAFNSTDATHLKFHPV
 Tetra 167 YEAWYGGYEMAAHQGRLTTHQORHAAAAAASFGFDOP---GFLHPHIVQRMASHSPVPE--PLAPLSAFCSTSISSSEGCY
 Fugu 168 YEAWYGGYPMTHAQHSRLSFHQORHAAAAA-----SFAFEQF---GFLHPHILQRMASHSPVPE--PLAPLSAFCSTSMSASSEGCY

Human 245 FVP-H-----SAPAVLTHNNESSNCVFEFHVHTPSSSGEGGTLIPQVYRNQVAVDNLQDELPL---PQSVGLHGXYLOPVTEEKH
 Mouse 245 FVP-H-----SAPAVLTHNNESSNCVFEFHAQTPSSS-CEGGGTLIPQVYRNQVAVDSNOEELS---PQSVGLHCHLOPVTEEQR
 Rat 245 FVP-H-----SAPAVLTHNNESSNCVFEFHAQTPSSSGEGGTLIPQVYRNQVAVDSNOEELS---PQSVGLHCHLOPVTEEQR
 Tetra 246 PPE-HAASPFAVSAPFVWADSAAPPTGNVFEFHL-----AAAGDPN-LMASRLFRARRSSXXXXXXXXXXXXXXXXXXXXX
 Fugu 247 PSEQHAAVSTFAISAPFVWMPD-AAPPTGNVFEFHL-----AAAGDPN-LMASRLFRARRSSMDLFIENDTGAGSGGSITYVRLQPVTEELY

Human 323 NYHAFELTVS-----VVEPIGQNWPIGSPHYSSDSSQITSSDPSDFQSPPTGAAAPFGSDVSMPEFHLPQTVLOESP---LFFCFPQG--
 Mouse 322 NNHAPELTIS-----VVEPMGOIWPICSPHYSSDSSQITSSILSDFQSPPTGGTAAPFGSDVSLPFIIRLPQTVLOESP---LFFCFPQG--
 Rat 323 NNHTPELTIS-----VVEPMQONWPVCSPEHYSSDSSQITSSNISDFQSPPTGGTAAPFGSDVSLPYIRLPQTVLOESP---LFFCFPQG--
 Tetra 330 XXXXXXXXXXXXXXXXXXXXXXXXXXXX--XXXXXXXEFQSPPEPP-LLPSFDSSCAQSIPIHSTATYDPEGGMEPIDPQGH
 Fugu 331 SYVSPPELPFPGSLFTHHAGLTAKDKSPDPSSDS---TASSDNGEFQSPPEPP-LLPSFDSSSAQSIPIHTASTAMYDPSTGMLPIDPQGH

Human 405 TTSQQLVLTAFSS-----GGSALHFOVIGKLPQLF*
 Mouse 404 TTSQQLVLSASYSS-----GGSTLHPQVIGKLSQFF*
 Rat 405 TTSQQLVLSASYSS-----GGSALHFOVIGKLSQFF*
 Tetra 417 PSMQSFPPPTPSSELPLGGA---PST--HLETLMQSAWARHGGVVPQAQPDMTYHESLIAMQAVKPSLVLLPF*
 Fugu 420 PSMQSFPPPTPSSELPLGGA---PST--QLETLMQSAWARHGGVVPQAQPDMTYHESLIAMQAVNPSLVLLRF*

Figure 3. Peptide sequence alignment of the human, mouse, rat, Tetraodon and Fugu sensorin orthologs.

```
Human: 10 LIQPQSMAPCGG-----TPTYPE-SQIFFPTIHERPVVSFPPTCPPKVAISQ 57
L Q S+ HP G + P+ + + FP+I ERP+SFSPPPT PPK +Q
Fugu: 13 LSQQHSVPHPYSGGAGGAGGNIPIASLPDPTTVLFPSIPERPIVSFPPTMPPKAYNNQ 72

Human: 58 RRKSTSFLEAQTHHFQPLLRVTVGQSLLPPGGSPTNWTPEAVVMLG---TTASRVTGESCE 114
RRKSTS IEA T HFQP G +L P G T P + M+ A++ G
Fugu: 73 RRKSTSILEAHTRHFQPAYTRYGSNLHPFSGMET-VEPPPFLMVNPGFAAAAQRLLGVGE 131

Human: 115 IQVHMFEPESQV-YSDYRPGVLPLEEAHYFIPOEAVYVAGVHYQARVAEQYEGIPYNSSV 173
+ + +PS Y D R E +A + HY+A Y G P +
Fugu: 132 LHLPQGSDPSLYGYKDMR--AEHGEMRRRLSLNQAAALLD--HYEAMA---YGGYPMTAHQ 184

Human: 174 LSSEPMKOIPEQKPVGQTSSSVFEFFSCQAFLVGLH-QNLRLDSDLGPGSPLSSISAPI 232
LS + Q+ +S FE P FL HL Q + S + P PL+ +SAP
Fugu: 185 LS---RLSFHQQRHAASAAAAAFQEOP---GFLHPLHIQRMSAHSVPVP--PLAPLSAPS 236

Human: 233 STDATRLKFHPVFVPH-----SAPAVLTHNNESRNCVFEFHVHTPSSSSGEGGI 283
T + P SAP V+ N VFEFH+ ++G+ +
Fugu: 237 GTSMSASSEGCYPSQPHAAVSTFAISAPPMPDAAPTGN-VFEFHL~----AAAGD-PNL 290

Human: 284 LPQRYRNQVAVDLNQEEELPPQSVGLHG---YLQPVTEEKHNYHAPELTVS-----VVPE 336
+ R++R R+ ++DL E+ G G LQPVTEE ++Y +PEL +
Fugu: 291 MASRLFRARRSSMDLPLEDNTGAGSGSGTYVRLQPVTEELYSYVSPLELPPGSLFLHH 350

Human: 337 IGQNWPIGSPEYSSDSQITSSDSPDFQSPPP 368
G SP+ SSDS + SSD +FQSPPP
Fugu: 351 AGLTAKDKSPDPSSDS--LASSDNGEFQSPPP 380
```

Fig. 5. Peptide sequence alignment of Fugu and Zebrafish orthologs of sensorin.

Fugu : 103 PPLEMVNPGFAAAQRLGVGEPLHLPGQSDPSLYGYKDMRAEHGEAMRRRLSLNQ 156
P +FM+ PG A QRL +PLHL ++ SLGYKD+RAE E +RRLSLNQ
Zfish: 3 PSMFMM-PGMPA--QRLT--DPLHLQPPTE-SLYGYKDVRAEQEETVRRRLSLNQ 146

[illegible]

Figure 6B-1

gtaagtaaatgcttaagagcatgtaataactacaatgagagccaatggatagctgctgctaaattaaaattctttgtacaaaaccaagtaacagtatgagtct
gaagtagaaaaataaaatgctgtgtctctaaagccttattatagagtaggatatagacattaatgtgtggcatttagctgaactgacctacatgaagca
gctatgtgcaagatttagattagctgcatcttactaaagggaacgctctctttagcacttaagatgtgaaaactattcacagcaacaaaagctggac
atacaaatgttttgcgggtgggcgagaaagagttttcaaaagcctgttaaaatgattactaagatttctgtaagattccctattgtacaaaacaca
ggtaagaccctcaaacattatagaatgctgcatgtgagtaactctagattgaggaatcctttatttgaatgactgttttaatttagcgcttgtgttgag
ttaatcatcagttgctgtgtgttctgccttttaaaagctgtgttgacacaagttcctcccaagtctcctccagtatatgttcttttgcttttcccttttagtt
gtaactgttttattactgatttatttcagaatcaacttatagaaattatttccctgaaacattttatgcttgatacacatacacacacaattatcc
tgtgttttacattagacactctgtttctcatatagagaatgttaagtagatcatcagataaataccccatcttttgttgacctcagaggaaaaataa
aatagctcatctgcctctgaaaggatttagaaaaataagtaactgtcactgtacagcactgttttagtgagatcctttaagcagacagattttcccttc
tcagaaagccttttaactaaactttgtggaattcttctttaaataagaaaactcttctctgttaagaaaataactctctgttgcgattttttat
cctacagcttggaataaaagatgttcattgtttcagagtcctgttaactcctgcttggctaggtctcaaaatgaatggccacagtaacacacataaaa
aactaatataactaaacttttaaaacaactctactttctgtcttttaaaataattgagggtagaataagaagaactttataaagactgcacagaaaaataatt
tctgtgttctgcacaaaataattactgtgcagcatttctgtcagtttctacctttaattgttttttctggattattataaactagattttgctaagactt
ttttttaaaatttctctgtctatactcatcatggggagcattaggtctctatatttagaaagcaagtagaatgacatgaagataattatttgagcca
ttaactagctatatttattgatgctctaaaagcagacgttcttgcttttaataagacagtattctatggaggtattatgaagaattattaaaattattgag
aatttaaaaaatttgtctaattctattctgtctcatttcagggaatgtggcttacaataatggacaaaattataagtgaaaatttagaaaattagaataa
aagaaataaggctagagagaaaatcaaaacaggaatgaaactaaaacaatcatatgatgataatgtataactctgttattgggtggccacagatttggccca
aagcattctacacagccaagttttggaaatgaaaacatagtttcatagttcacacaatctaaaaaggaaaataaacagtccttattcaagacacacctttatg
gtattgatcatactgccttatagccatattaaacccctgtgcctgaacatttttttaagttttttttttttttttgacagcgtctcacctctgtctccca
ggtgggaaacacagtggtgatcttggctcactgcaacttcaacctcctgggtcaagtgttctcctgcctcagcctcccaagtagctaagattacag
gcaccatcacccagccagctaaattttgtatttttttagtagagacacagtttccaccgtgttggttaggtgtgaaactcttgaaactcttgacctcaagt
atccacctgcctggcctcccaagtgccaggattacaggtgtgagccacccatgccggcctaaaattattatttagagattcagattgaagatatctc
ccattgattactttgtaaaagtacaaaaattttcctatatccagttgtgaatttctctagaaaagctaaaacttttccctagtagatcatgactttttta
cctttacaaaataaagagtggaagtataaagaaatactgtgagccaaactttccatctcatgtcttctaaaaatagaaaatgagaatcctttttctgaggt
ttccagagctcttgcaaaaataatccaatcttaaaattctctgtatcttttaagtactttaaactaatagaattcattctgttgtaactcttaagtttgc
caagttgcaagctaataattcattgttttccctgtctgcaccttcagtcagttggtcattctttaaattcatcacactgttaacacttggtgttatttca
gaatattagcatgataaactgatcattatcatgatcattcaaaacctgtccttgactgacaacgcagaatagttgtgaatagttgtgatctgtcaga
tgtgctatgttgagttatagctctcctcaaaaaaaaactgactttgtggaattgggagaggatggaacatttcttcatgcaactatgccctgttatca

9/22

ttgtacttttgtttattttttcacagtagctgtgtttttcatgtgtgtgtgtttgtttgtgtgtgagcctcgtcgtgagcctcgtgagccttagcatgtcggtttgtgtgt
tcccacttttctgctgttgcctctgtgtccgcgcatctctccagtgctcttccacccacccagccagtagtctgctgcacctcttttctccttccctcctcc
ggactgccccgaggaaaacttttggcggaaaagcttttctaagcatggagagtgtcctgcctatgactctgctctcagcgcaagcacgcagcgtccag
cctgccttccctctttgtcagtagctgtatgttaactgtaaacttctgacaaatgaacaattattaccaatgaatacatccaggcaacaagaattttaatt
aaaattgaataaagaacaggactaaacttggttataattatgcttctaggtataccagttattcctgtaggaacttttttgaataaggctttaaggggg
aataaaaaagcaaggaggtaatgagagagaatacctataatgtccttggcacataccagcatgtgtccaataaaaatttacccttccattatattcctgt
ttcctatcgacataaaaattcttcattttcataattctagtttaaaaggattaaattttaaagttattgcttaatacaaaagcaaaattttatcaaacat
acccaaacatttttctcctcactttattctaaaggagacactaaataaaactggatgagtcctttttttttaataaaccttagtcagcaaaagaataaacc
ttggagtgaacatttttaatttaatacatttatgtaaagtgtatcagcaaaccaagaactacataatgttcccccatggaagtatttttctcct
taggtttcaacctaaagggtgagctgtagctctaataatgaatgcattagatgcataatagtatgactttcttggctccattaatgtttataaaaaga
tctatcttataaaaattattttaagcatgtgggtccttgcgtctttccaaaataaagaatatagttaaactttctgaactagcaacccagaatttttcaga
ttttacttaggagaatgaagaactatacagtggtatttatggctttgagacataccatgagtaaaatacaataaaaatgagcatctgactgaaaaataga
aatccttatttagctattgtgaacagttgatctactttgacagctcgataaggatcttccctatggaaagaatgaaactgaaactcttttaaaaagcgggg
gagcgtgtgtagtgaattcttactattattacttattttatgttttggagatggaaactcgcctgtgtgccaggctggagtgcagtggtggcagcat
ctcgactcaggcaacctctgcctcctgggttcaagtgattctcctgcctcagcctccgagtagctaggactacagggcgcatgccacacacccagct
aattttgtatttagtagagatgggtttccaccatgtttggccaggatggctctcctgacattgtgtatctcctgacattgtgatccaccccttcagcctccacagtcg
tgggattacaggcgtgagccacgcgcgggtaataatgaattattatagtgatattgttgcatttggcaagctgtgttctctaaatttagtcttaa
atctagaactgcttctagaataaaaagtgcatacaaaataaaaattggctgtcatttatattacccttttcagttgattagaacaccaggtgtcagaatc
agaaaattcatggttacggccgggtcagtggtcagcgcctgttaattcctagcactttgggagggccgaagtgcagagatcaactgaggtcaggagttcga
gaccagcctggcccaacatggcgaaatgtgtctctactaaaaatacaaaaaattagctgggCGTGGTGGCGCACGCGTGTAGTCCAGCTACTTGGGAGGC
TGAGGCGAGAAGATCCCTTGAAACCCAGGAGGGGAGTTGCAGTAAGCCAAGATTGTCTACTCCAGCTCCAGCTGGGCGAAGAGAGAACTCCATCTCA
AAAAAAGAAAAAGAAATTTTCATGGTTATGCAACTCTTAATTATGATCAGAAAAATGGACATTTTGTGATTTAACTCTGTAAACATGTTTCATGTAGTA
AAAAATAATAAAACTATTAAATCATCTAGCTTGGGAGAGATAGGAGAAAGACATTACTGTCTCACTAGTCAAAATTAATATCTTTTACTATCCACCAAAAT
TCTCTTCTGATTTCTGGTTAGGAAGGCATACTATTAAATTGATAAGAAAAATAAACTGAAGGCCTCTAAACATATCACAGGTAATAAAGAAATATAGGGAAG
TTAGTTCAATAGTTTAAATTAAAGCACACTTCTTACAGTATAGAACTAGTCGGGCTTTTATGCTTGTATTAGTCTTACTCTTCTTAACTCTTTTT
CTGTTGATGTAATTACATTAAATGCTTAAAGTGAACTTTTAAAGTGTGGGTAAAAACGAATAATTACTTACAAAAGTTTAAATCTTCCATTTCCCTTTG
AGAGAGGAAAGTTATGGAAGCAGCTCTTATCTAAAGCAAAGACCCACAGATTGATTTTCATTTGGCCCTGGATGTATTAATGGATTTTACTATGCA
CATAAATTTCCAGAAAGCATTTTATTATTATTAATAATAAATTTAGTGTAAACCATTTTCATAGGGTTACACAGAACTACCCAGTTGTGCATGCTCTGATG
TAATTTTCACATATGAATG

Figure 6C

[illegible]

Figure 6D

gcagcttttcttccattcaaggaagtccacagattgactttattggccctggatgtattaatggatttttggctataaaattcc
aaaagcatgttagttattaaaaatttgggtgtagccacttcattgggtttacacagaaactatccagctgtgcatgtctgtatgtaaat
tttacatatgtgaATGATGAATTAATTCATGTTGATACAGCCTCAGTCCGTGGCGCATCCGTGTGGGGAGACCCCAA
CATACCCAGAATCACAGATATTTTCCCAACTATTTCATGAACGTCCAGTTTCTTTTACCACTCTACCTGTCACCCGAAAGT
AGCCATTTCCAGCGACGTAAGAGTACCTCCTTCCCTGGAAAGCTCAAACCTCGCCACTTCCAACCCCTGCTGAGGACTGTTGGCCAA
AATCATCTTCCACCTGGTGGCAGCCCAACTAACTGGACACCAGAGGCCATAGTTATGTTGGTACTACAGCCAATAGAGTAAATA
GAGAGCTATGTGAGATGCAGGTCCAACCTGTGTTTGAGACAACCCAGATTTACAGTGACTATAGACCTGGACTAGTACTGGCAGA
AGAAGCTCGCTATTTTATCCTCAGGGAACAGTATATCTAGTGGGTGCATTACCATCCCCATCGGCAGGACAGTATGAGGGT
ATTTTCATACAACCTCACCCAGTACTGTCAAGTCCATGAAAACAGATAAAGTGAACAGCCAGTGCCTGGTGCCTGCCTCAAGTT
CTGTCTTTGAATTTCTTGGACAGGCTTTCTAGTAGGACACCTTCAGAAATTTAAGATTAGATTCTGGACCAAGTCCAGCATC
ACCCCTGTCTAGTATTTCTGCGCCTAACAGTACAGATGCTACACATTTGAAATTTCAACCTGTCTTTGTACCTCATTCTGCACCA
GCTGTGTTAACTCATAGCAATGAGAACAGAACTGTGTATTTGAATTTTCATGCTCAAAACACCAAGTTCTCTTCAGGAGAAG
GAGGTGGGATTTTACCTCAACGTGTTTACCGAAATCGACAGGTTGCAGTGGACTCAAGTCAAGGAGAACTGTCTCCTCAGTCAGT
TGGATTACACTGCCACCTGCAGCCTGTTACTGAAGAACAGCGTAATAACCATACCCAGAAATTGACCAATTTCTGTGGTAGAACCT
ATGGGACAAAACCTGGCCAGTAGGAAGCCAGAAATATCCAGTGATTCCTCTCAAATTAATCTTCAGATATCAGTGATTTTCAAT
CACCTCCCTACAGGGGAACAGCTGCACCTTTTGGCTCTGACGTCTCATTTGCCCTATATTCGCCCTGCCTCAGACAGTGTACCA
AGAATCCCCACTATTCTTCTGTTTCCCCCAAGGAACCATCTCAGCAGGTCTTATCTGCTTCATATTCTTCAGGAGGATCTGCA
CTCCATCCACAGGTTATAGGAAAACCTTTCACAAATTTTAAaataccctattatgcaccaatgcatttaaatctctgtctcat
ttccctgaatcatggatttataaagaaatgtatttttcttagtaggatttctttaggttataaataatgtactaaaaaaatga
actttaaagtggaaaaataacctgggtttacctgtgatttttagttgtcagaatcaatgactaatatctgtggaatttttgggttagtag
taaaaaatgcaattgtc

Figure 6E-1

ccctcaagtatctacaaagatggtagggaatgcataatgcagattttacagtttcggttaagccagtcctgagagaattcttcttctctcc
 tcttctacttttttacaatgccttttttttttttggcggggtggtggtgggggtgtgttttcttcaataactactgcttaatt
 taccttttattctgtagtagagtcctgggtatgtctgtgaagtgatcacaaagaccatggctaaagctatcaaaagacagagtagcattaat
 taagaggaaacgagagcagcggcagttggtacgggaggaagcaaaaaaaagcaggaagagagcagtcctcaaacagcaggtagaacaat
 ccagtgcctccagacaggaatcaagcagtcctctgtgtagcaccggcatacctactgcttctaccacttcagcttcagttctacacat
 gtagaacctgaagaacctgaggcagatcaacatcaacaactacagtaaccagcaaccagtatatctgtgttatgtacgtatcttgggaagt
 gacagataggctaattgcctcaattgggttttttagcttctgttatttgcactggcttttgggttttgaattatgaatcctactatt
 atacaaaatttgaaatggttgcgtctagcttctcttattcttatactgtgactaggagaatagaaggatgggggtggttaaggaaaatgtga
 ttaggatacaaaaataatttataagcttcaacatcttgagggtttcaattttagatttctactttcacaaatttatataaaaaacagat
 ttaataatgaagaatgaacaagcttggctactaaaaattgaaccatcagtcctggttgagctcttttagcaaaagctgaatggtatatagacta
 aaaactaatgagaacactggttatttaattacttgtttaatactgagaacctcctggcctccataaattaagactttgaaagctgtgtc
 aaatttaaggcttataggtttcatgctgtttcccttactcctcaggtggttaagattctgtgtagccatatggattcctcatttaaatata
 gaaatctactgaatatggctacaattcatttttatttagttcaaatctgtgcaagggaataaactagttaccctaaataataatgggtctaaat
 atgagaaaatgtgaacctccattttgtgattgtcttctctctctctctctctctctctctctctctctctctctctctctctctcag

XH03x08 (PRKWNK1)

CTGATGGGACGGTTGACAGTGGTCAGGGATCCTCTGTCTTCACAGAATCTCGAGTGAGCGCAACACAGACAGTTTCATATGGTCCCAACAT
GAACAGGCACATTCTACAGGCACAGTCCCAGGGCATATACCTTCTACTGTCTCAAGCACAGTCTCAGCCCCATGGGGTATATCCACCCCTCAAG
TGTG

gtaagtaaatgcttaagagcatgtaatactacaatgagagccaaatggatagtcgtctaaattaaaaattcttgtacaaaccaagtaaacagta
 tgagtctgaagttagaaaataataaaatgctgtgtctctaaagccttattatagagtaggatatagacattaatgtgtgggcattagctgaact
 gcctacatgaagcagctatgtgcaagatttagattagctgtcattttactaaaggaacgctctctttagcacttaagatgttgaaactat
 tcacagcaaaaaagctggacatacaaaaatgtttgcggtgggagagagttttcaaaagctgttataaatgattactaagtatctcc
 tgtaagattccctattgtacaaaaacacaggttaagaccttcaaacattatagaatgctgcatgtgagtaactctagattgagggaaatccttta
 tttgaatgactgttttaatttagcgcttgtgttgagttaatcatcagttgctgtgtgttctccttttaaagtcgtgttggacacaagt
 tccttccccagtatatgttcttttgccttttcccttttagttggtaactgttttattactgatttatttccagaatcaacttatagaaaattat
 ttcttgaaaacattttatgcttgatacacacacacacaattatcctgtgttttacattagacactcttgtttctcatatagagaatgt

T/CA2003/001716

Figure 6E-2

taaagtagatcatcataaccatttctttgttgacctcagaggaaaaataaaatacgtcatctgcctctgaaaggtatttagaaaaata
 agtaacttgtcactgtacagcacttgttttagtgagatctttaagcagacagattttcccttctcagaaagcttttaactctaaactttgtggtg
 aattcttgctctctttaaatgaaaatctttctctgttaagaaaatacttctgctgttgctgatttttatcctacagcttggaataaaaagat
 gttcatttgtttcagagtcggtttaatcctgctttggctaggtctcaaaatgaatggccacagtaacacaaaaataaaactaatatactaac
 tttttaaaacaatctacttctgtctttaaaataattgaggttagaataagaagaactttataaagactgcacagaaaaatatattctgtgtt
 ctgacaaaataattacttgtcagcatcttctgtcagtttctacctttaattgttttttctggtattataaactagattttgtctaagactt
 tttttaaaaatttctctgtctatactcatcaggggagcataggctctctatttagaaagcaagtagaatgacatgaagatatatttt
 gagccattaatctagctataatttattgatgctctaaaagcagacgttcttgcttttaatagcagtatctctatggaggtattatgaagaatt
 attaaaatttgagaatttaaaaaattgtctaattctattctgtctcatttcagggaattggcttacaataatggacaaaattataagtga
 aatttagaaaattagaataaaagaaataaaggctagagagaaatacaaaacagggaatgaaactaaaacatatgatgatataataacttg
 gtattgggtggccacagatttggcccaagcattctacacagccaaagttttggaatgaaaacatatagtttcatagttccacaacatctaaag
 gaaaaataacagtccttattcaagacaacctttatggtattgatcactgccttatagccatatataaacctgtgcctgaacatttttttaag
 ttttttttttttttgacagcgtctcactctgtctcccaggtgggaatacagtggtggtgactcttggtcactgcacactttcacctcctgg
 gttcaagtgattctcctgcctcagcctcccagtagtaagattacaggccacccatcacccagccagctaaattttgtgatttttttagta
 gagacacagtttcaaccgttggctaggtgttgaaactcttgacctcaagtgatccacctgcctcggtcccaaaagtccaggattac
 aggtgtgagccaccatgcccgccctaaattattatttagagattcagattgaagataattctccattgattactttgtaaaagtacaaaaata
 tttcctatatccagttgtgaatttctctagaaaagctaaacttttcttagtagatcatgactttttacctttacaaaataaagagtggtgcaa
 gtataagaaaatactgtgagccaaacactttccatctcatgtcttcttaaaaatagaaaatgagaatcttttctgaggtttccagagctcttgca
 aaataatccaatcttaaaattctctgtatcttttaattgtactttaaaactaataagaattcattcttgtgaatcttaagtttgccaaagtgc
 aagctaataattcattgtttccctgtctgcaccttcagtcagttggtcattcttaaaattcatacacttgtaacacttgggtgtatttcaga
 atatttagcatgataactgatcattatcatgatcattcattaaacctgtcctgactgacaaacgcagaatagttgtgaatagttgtgatctg
 tcagatgtgctatgttgagttatagctctcctcaaaaaaaactgactttgtggaattgggagaggtggaacatttctcatgcaactat
 gccctgttatcatattgtactttgtttattgtttcacagtaactgtgttttcatgtgtgtgtttgtttgtgtgagcctcgtcgtggcc
 gtagcatgtcgggttgttccatctttctgtgttgcctgtgtcccgcatctctccagtgctcttccaccccccagcctactgtc
 tgcacctcttctcctcctccgactgccccgaggaacttttgcgaaaaagctttctaagcattggagagtgctcctgctatgca
 ctctgctctcagcgcaagcagcgtccagcctgccttccctcttctgtcagtaactgtatgtaaacttctgacaaaatgaacaat
 tattaccaatgaatacatccaggcaacaagaattttaattaaaattgaataaaagacaggaactaaacttgggtatatattatgcttctaggata
 ccagttattcctgtaggaaaactttttgttaataaggctttaagggggaataaaagcaaggaggtaatgagagagaataacctataatgtcct

T/CA2003/001716

tggcacataccagcatgtgtccaataaaaatttaccttttcattatatacttctgttttccctatcgacataaaaaattcttttcattttccatatctcta
gttttaaaaggattaaatttaaagtattgccttaatacaaaagcaaaattttatcaaacattaccacaacattttttctcctcactttatttct
aaaggagacactaataaaaactggatgagtccttttttttaataaccttagtcagcaaaagataaaccttggagtgaacattttattttaatt
taatcatttatgtaagatgatgtatcagcaaaccaagaactacataatgttcccccatggaaagtatttttcccttaggtttcaaacctaagg
tgggactgtagctctaataatgaatgcattagatgcataatagatgtatgactttcttggctccattaatgtttataaaaagatctatcttat
aaaaattattttaagcatgtgggtccttggctgttttccaaataaaagaatatagttaactttctgaactagcaaccccgagaattttcagatt
ttacttaggagaatgaagaactatacaggttatattatggctttgagacataccatgagtaaaataacaataaaaatgagcatcttgactgaaaa
tagaaatccttattagctattgtgaacagttgatctactttgccagctcgataaggatcttccctatggaaagaatgaaactgaattctttt
aaaaagcggggaggctgtgatatgtgaattcttactattattacttattttatgttttgagatggaaatctcgtctgtcgcaggct
ggagtgcagtggcacgactctcgactcacggcacaacctctgcctcctgggttcaagtgtattctcctgcctcagcctcccagtagctaggacta
caggcgcatgccaccacacacccagctaaatttttgtatttttagtagagatgggtttcacccatgtttggccaggatgggtcttgatctcctgacat
tgtgatccacccccctcagcctcccacagtgctgggattacagggcgtgagccacccgcgcagcgggtaataatgaattatatagtgatatattg
ttttgcatttggcaagctgtgtttctaaatttagtcttaaatctagaactgcttctagaataaaaaagtgctatacaataaaaattggctgtcat
ttatatattaccccttttcagtttgattagaacacacaggatgtcagaatacagaaaaattcatggttacggccgggctcagtgggctcacgcctgtta
cctagcacctttgggagggccgaagtgagcagatcacctgaggtcaggttcgagaccagcctggccaacatggcgaaaatgtgtctctactaa
aaatacaaaaaattagctggg

CGTGGTGGCGCACGGGTGTAGTCCAGCTACTTGGGAGGCTGAGGCAGAAAGAAATCCCTTGAACCCAGGAGGCGGAGGTGCAAGCAAG
ATTGTCACTGCACCTCCAGCCTGGCGGAAAGAGGAAAGACTCCACTTCACAAAAAAGAAAAAATTTTCATGGTTATGCAACTCTTATTAT
GATCAGAAAAAATGGACATTTTGTGATTTAACTCTGTAACATGTTTCATGTAGTAAAAAATATAAAAACTATTAAATCATCTAGCTTGGGAGA
GATAGGAGAAAAAGACATTACTGTCTCACTAGTCAAAATTATATATCTTTTACTATCCACCAAAAAATCTCTTCTGTGATTTCTGGTTAGAAAGGCATACT
ATTAAATTGATAAGAAAAATAAACTGAAGGCCTCTAACATATCACAGGGTAATAAGAAATATAGGGAAGTTAGTTCAAATAGTTTAAATTAAAG
CACACTTCTTACAGTATAGAACTAGTCGGGCTTTTATGCTTGTGTTTAACTCTCTCTTAACTCTTTTCTGTGTGATGTAATTTTAC
ATTAAATTAAATGCTTAAAGAGTGAACTTTTAAAGTGTGGGTAAAAACGAAATAATTACTTACAAAGTTTAAATCTTCCATTTCTTTGAGAGAGGAAA
GTTATGGAAAAAGCAGCTCTTATCTAAAGCAAAGAGCCCCACAGATTGATTTTCATTTGGCCCTGGATGTATTTAATGGATTTTACTTACTATGCACAT

Figure 6E-4

AATTTCCAGAAGCATTTGTTAATTTAATTAATAAATTTAGTGTAACCATTTTCATAGGGTTACACAGAACTACCCAGTTGTGTCATGTCT
 GATGTAATTTTCACATATGAATGTATGAATTTACTTGTCTTTATTCATGTTGATACAGCCTCAGTCCATGGCGCATCCGTGTGGGGACCCCAA
 CATACCCAGAATCACAGATATTTTCCCAACTAATTCATGAACGTCCAGTTTCTTTTTCACACCTCCCACTGCCCCACCGAAAGTAGCCATT
 TCCCAGCGCGTAAGAGCACCTCCTTCTGGAAGCCCAAACTCACCACTTCCAAACCCCTGCTGAGGACTGTTGGCCAAAGTCTTCTTCCACC
 TGGTGGCAGCCCCAATACTGGGACACCAAGAGGCCGTAGTTATGTTGGGTACTACAGCCAGTAGAGTAACCTGGAGAGTCATGTGAGATACAGG
 TCCATCCTATGTTTGAACCATCTCAAGTTTACAGTGACTAGTACTTCCAGAGAAGCTCACTATTTTATTTCCCTCAGGA
 GCAGTGTATGTAGCTGGGTACATTACAGGCCCGGGTGGCAGAACAGTATGAGGGCATTTCCATACAACTCATCAGTACTGTCAAGTCCCTAT
 GAAACAGATACCTGAACAGAACCCAGTACAAGGGGCCCTACTTCAAGTTCTGTCTTTGAATTTCCATCTGGAACAGGCTTTTCCGTGGTAGGAC
 ACCTTCAGAAATCTAAGATTAGATTCTGGATTGGGTCCGGGATCTCCCTCTCTAGTATTTCTGCACTATTCAGTACAGATGCTACACGTTTG
 AAATTTCAACCTGTCTTGTCTCTCATTTCTGCGCTGTGTGTTAACTCATAAACAATGAGAGCAGAAAGCAACTGTGTATTTTGAATTTTCATGT
 TCACACACCAAGCTCCTCTTTCAGGAGAAAGGAGGTGGAATTTTACCTCAGCGTGTTTACCGAAATCGGGCAGGTTGCAGTGGACTTGAATCAAG
 AAGAACTGCCCTCCTCAATCAGTTGGATTACATGGCTACTTGCAGCCTGTGACTGAAGAAAGCATAAATACCATGCCCCAGAAATTGACCGTT
 TCTGTGGTAGAGCCTATCGGACAGAACTGGCCCAATAGGAAGCCAGAAATATTCAGTGATTCCTCACAAATCACTTCTTCAGACCCCACTGA
 TTTTTCAGTCACTCCCTACAGGGGAGCAGCTGCACCTTTTGGCTCTGACGTCCTCAATGCCCCCTTTATCCATCTGCCCTCAGACAGTGTAC
 AAGAAATCCCCACTTTTCTTCTGTCTTCCCCCAAGGAACCACTCTCAGCAGGCTCTTAACTGCCCTCATTTTCTTCAGGAGGATCTGCACCTTCAT
 CCACAGGTTATAGGAAAACCTTCCACAATTAATTTTAAACTACCTTCTGCACTTAAACAATTTTCTATTTCTTATTTCCCTGAATC
 ATGGAATTTGGAGAAATATTTGTTTAAATTTTATCAGTAGAGTTTCCCATCTTTTGGGGGTTGTGAACATACATATATGCAATTTAAACAAAAA
 TGTGAGAGAAAGCTACCTGATTTTACCTATTATATGTGAAAACCCAGTGGAAAACACAAAACTAGAAATTTTAGTCACTTCTTCACAAATCACGA
 CTTCTATGCACGTTATTTTCAACCAAGTAGTGAATAATGCAAGTGTATGTAATGTATGGTTGACCCAGCATTAATTTAGGAATACAAATCTTTAAG
 TATTACTTTTCTTCTCCCAACAAGTTTTTAAATAAATAGGATAAAATTTTCTATAAAATATAAAACATGGAAAATAGGGAATGCTGTTTT
 TGAGGTAATAATTAATAACACAGAAATTTTCATTAGTGTGGAAGGATCTAAATAAGACAAAAGTATATCATGGGAATAAAAAAGATAGAAAAG
 GAAACAGTTTAGGAATTTGCTTAAACAAATGAAAATGCCCTTTTAAATAATGGCATCAGTCAAGCAAGTTGCTGTGCAATTAATATATGTCCAAA
 TAAAAATGCTAATTCATAAAATTAAGG

Figure 6E-5

actgattctaatTTTTtagcaactaacaatctagtgggtgggctTTTTTTTgttataagaattTgtgatatgttgcaagctaaggaa
 cagtagacatcttaggccttgatttctcaacaaatttaatttagttatttatgtgtatgagcatatcacacatttagcaaataaagaca
 atgtaatgaatgtaataaattcaaaagtacagtcagtggttactgtttcattatattcatttggtatatttatcaaatgtgtacatcttat
 aaaatctacatagaagactatgcagttatattgaagacattTggatatcctacataaagtactatgtgtgttcttctactttcttcc
 atgttcttccctcatcaggagcataaaacactcccaaaatgtgcacaaattatcaattatttagtgcgtactaaagacacaaaaagatccag
 aagcttgacagctatacatgggcagtcagtataattcacatcagattcaagttgcagccctattataaatttaattgattgccagtaatt
 atcctttagactttccaaagtataaaataaccaggtagttaattatagacatactacagaaggaaatcaatattataatagttaagctgcag
 tggaaTgtgatattatatccatgttaagccttaaggaaacccctgtaaactgaaatggaatgacaaaagtggtagcccttttcatttcgttaac
 atacctatataatagtcctgcattctcactttctagtgaacacactgtagttgggtttctcagactcaacactgttgacatctgagg
 ttggataaattcttTgtgtgggtggttctctatgttaggctatttagaacatccctgggtttatacctcctaaatactagtagcagca
 atctccctcagttgtgacaaaccagaaatgtctccacataattgccaaagtccccctgggggtggcaggaattgctcctggatagcaatgatcg
 ccotaggtaactggcagttgatgataacagatttttgatacttttgataaaagtcatgtggcactacaagaagtattaccaaaattagcat
 taccaaatacccttcttaattcttcagataaaatttcagattgatagattataactttgttccctgattattgtatttctataaacaattt
 tattatgaattctgacgattctagccagctaaaaattacagcctgagcagcattgacagcttaggttttcactgcagtggttttgcagtgggt
 tgggatcagaacatgttggtggtatgatcccttttctaaaaattttctacaaaagtTgacccagaggcctctcccatataaatcaggttaa
 tgtttcatTTTgatcaagcaaacctctgaggagattaaatatattcattTgcáaaagcctgacctctatacacactttaggccttctctaat
 gttgtgtbcacttcttcttcag

XH03x09 (PRKWNK1)

GCACAGGGGCAGAGCCAGGGTCAGCCATCCTCAAGTAGCTTAACAGGGGTTTCATCTTCCCAACCCATACACATCCTCAGCAG

T/CA2003/001716

gtgagaacaaatgcattgcaacattttatgaattagcagtgccagaacactatttttaacttttTgtgattTgttatataaatggcctgctaag
 tttttgaaaggaaaaattaaagaggcataccatgtcttggttcagtaacaaagaaaaagagacttttctgtgtttcagaaaaatctcttt
 tgtttaataatacttttgaacaaatttagtagacatacatattactgatataaattatcttttaaaagaggattctcacacctttttttgaga
 caaagtctcactctgtcaccaggctggagtgagtggtgcaattcttagctcactgcaaccccgctccagggtcaagcagtcctcctgc
 ctacagctttccacgtagctgggactacaggtgcctgccaccacactcagctaaattttTgtatttttagtagagacgggtttcaccatgctg
 gccaggctggtctcaaacctcctgacctcaggttaattccacccctcctcggttcccaaaagtctgggattacagggcatgcgccacctcaaccag

Figure 6E-6

cctggattctacacataatttaacgggtacaaagagatgtcattccttttctctctttttccactgcttgcagcagtgatttgtcattacctt
 cattgtgcttgccctactaatacctgcacttcttttaattattctaatctgtagaatgatgtgcagattttacgctaagaagcatcttaa
 tgaacacctaagaacaattacttcaagaacatttatttatttatttatttattttagacagccttgctctgttggccaggctgga
 gtgcagtggtgcaatcttggtcactgcaacttccatctccagggtcaagctattctcagcctcagccgcccagtagctgggactacag
 gcggtgccactatgcccagctaaattttgtattacgttctacttatgctggctaattagtagagatggcggttccacatggtggccagg
 ctggtcttgaaactcctgacctcaagtgatccacccgcttgccctcccaagtgctgggagccactgtgccagcccaaatattttaactt
 tgtttgggggctgttccagaaagtccaaggaattcaagagattagtgctcatcctgatacatccaacagtaaaactggtaattctaaat
 taaatattatagtgcatcactgataatagttaatagtaattattgcaagggaataatgcatttttggcataatgaatacttattagcta
 ttctacataaaattttagccactatgtatgtaatgcaaaacttctaatgttagtccccatgtgggttttttaatacagcaaaactctaaact
 ggcttctgttctattagtttgggttggaaacaaagttttcggttcttaattacaaatggcccttttccagtcacaacatatggtcagtga
 gtcatctcgatactggtttaagatcaggaggtgctgtacaagatgaaaaactcactaaaggcactttctacagaaacagtgattgattgtt
 tatacatgcattcactccgtgaactgaaggactgggacttatatgctttgtttgcacctgtccttttccaaacaaagactgggttaaattc
 aagtaaacataattttgccaaccttaacaaaactgttttattctattacaggttccataatagaatgcaaatcacagtcacataaagctg
 gttgagaatggaagtactgagtactaatctcctaactgcttatcttggatgtagacgaacaccaaatagaactgagcaaatgtatgtgctc
 tgtttttctttccaaagggaattggttatctgtctgtatttgggatatgtatgtatacttactatactcaaatgaagaagataataaataa
 ttcataggttcatgttaagaaagaggtgtgagaaagagagagcagtaaccataaaatgtcagattaaaagataggtagctacctctcca
 ggagaaaaaataatggctaaatgctatctctacagcagaaacaaaataatatttgtttcagtgatgtttttagatacttttag
 catgtgtgtagaaatcacaaaattgagccctcgatatataataggaagtgtattcgggtctagaataattaaacttctccaaagtcacttagata
 ataaaggagctacaaatataaccagaaattcaagattcctagaatagtattgtttcttccattttaaataaagtctgcagggacttaagt
 attctgcaaatgtacctcagaagctgtctttattgttttaattattggaggtccacgtttgaaaaaagaaagggtgaaaaactttttgtact
 atatggatagtctatttttagacatgattttgttttacgtactttcagggtcccaatccacacttagagaaccatatctctgttttatttt
 agtgaggaagaaataacttttgcataaaatttactttattctatgctgtttcttgaaacatagtaaatgaggaattgctttttgtgtgtttt
 tgtagggtctcgatagtgatatactagatcttagcaagggtaccaaaaatagctgcctcctgctgaaaaagcttgtgaccacaaaacaaagt
 gtgaactttgtcagacacagtgccagctacttggaaaggctgaggcaggaagatcaacttgagccaggaggttcaagggtacagtgaaactatga
 tattggcaccatactccagcctgggcaacagagcatgactcatctcaaaaaaaaaaaaaagcaccggaagtgtgaactacctacccctactatcc
 ataaagtctgtttgaactggtttcatgaggcaagcttgggcagagcatgatgctaaactgttaagctagaccacccctgcctggctgaagt
 ctgattattagtcagtgaggactgcgcacagcagatcatgctacttctggcatatccagaggt

cgtggtggcgacgcgtgtagtccagctacttgggaggtgagggcaggaatcccttgaacccagggagggttcagtaagccaagattgtc
actgcactccagcctggcgaaagaggaactccatctcaaaaaaagaaaaaattcatggtatgcaactcttatttatgatcagaaaaat
ggacatttgtgatttaactctgtaacatggttcattagtaaaaaataataaaactattaatcatctagcttgggagagataggagaaagacatta
ctgtcactagtcaaatatatactctttactatccaccaaaaatctctctgattctgtgttagaaggcatactattaatgataagaaaaataaaact
gaaggcctctaacatatcacagggtaataagaatatagggaagttagttcaatagtttaaataaagcacactcttaccagtagaactagtccggg
cttttatgacctgtttagttcttactcttctttaactcttttctgttgatgtaatttacattaatgcttaagagtgaactttttaagtgtgggta
aaaaagaaataaattactacaagtttaattcttccatttcccttgagagaggaagttagtgaagcagctcttatctaagcaagagagccacag
attgatttcattggccctggatgtatttaattggaattttactatgcacataatttccagaagcattggtattattattataataatttagtgta
accatttcatagggttacacagaactaccagttgtgcatgtctgatgtaattcacatatgaATGTATGAATTACTTGTCTTATTTCATGTTGATACA
GCCTCAGTCCATGGCGCATCCGTGGGGGACCCCAACATACCCAGAATCACAGATAATTTTCCCAACTATTCATGAACGTCACGTTTCTTTTCAC
CACCTCCCACCTGCCCAACCGAAAGTAGCCATTTCCCAGCGGGTAAGAGCACCTCCTTCTCGGAAGCCCAAACTCACCACTTCCAACCCCTGCTGAGG
ACTGTTGGCCAAAGTCTTCTCCACCTGGTGGCAGCCCAACTAACTGGACACAGAGCCGTAGTTATGTTGGGTACTACAGCCAGTAGAGTAACCTGG
AGAGTCATGTGAGATACAGGTCATCCTATGTTTGAACCATCTCAAGTTTACAGTGACTATAGACCTGGACTAGTACTTCCAGAAGAAGCTCACATATT
TTATTCTCTCAGGAAGCAGTGTATGTAGCTGGGTACATTTACCAGGCCCGGTGGCAGAACAGTATGAGGGCATTCCTACATACAACTCATCAGTACTGTCA
AGTCCCTATGAACAGATACCTGAACAGAGCCAGTACAAGGGGGCCCTACTTCAAGTTCTGTCTTTGATTTCCATCTGGACAGGCTTTCCTGGTAGGA
CACCTTCAGAATCTAAGATTAGATTCTGGATTGGGTCCGGATCTCCCTCTCTAGTATTTCTGCACCTATCAGTACAGATGCTACACGTTTGAAATTT
TCACCCCTGTCTTTGTTCTTCATCTTCGCCCTGTGTGTTAACTCATAACAATGAGAGCAGAACAACTGTGTATTTGAATTTTCATGTTTCACACACCAA
GCTCCTCTTCAGGAGAGGAGGTGAATTTTACCTCAGCGTGTTTACCGAAATCGGCAGGTTCGAGTGGACTTGAATCAAGAAGAACTGCCTCCTCAA
TCAGTTGGATTACATGGCTACTTGCAGCCTGTGACTGAAGAAAAGCATAATTACCATGCCCAAGAAATTGACCGTTTCTGTGGTAGAGCCTATCGGACA
GAACTGGCCCAATAGGAAGCCCAGAAATATCCAGTGATTCCCTCACAAAATCACTTCTTCAGACCCCAAGTGATTTTCAGTCACCTCCCCCTACAGGGGGAG
CAGCTGCACCTTTTGGCTCTGACGTCTCAATGCCCTTTATCCATCTGCCCTCAGACAGTGTACAGAATCCCCACTTTTCTTCTGTTCCCCCAAGGA
ACCACATCTCAGCAGGTCTTAACTGCCCTCATTTTCTTCAGGAGGATCTGCACCTTCATCCACAGGTTATAGGAAAACCTCCACAAATTATTTTAAactac
cctactttgcaccataaacatttaaaatttctatttcccttatatttccctgaatcatggattttggagaaaattgtttaattttatcagtagagtttcccc
atctttgggggtgtgaactacatatatgcatttaaaaaaataatgtgagagaagctacctgatttacctattatgtgaaaaaccagtggaaaaaa
cacaataaactagaatttttagtcattcttcacaatcacgacttctatgcaggttattttcaaccagtagtgaaaatgcaagtgtatgtaatgtatggtt
gacccagcattatttaggaatacaaatcttaagtattactttcttccaaaacagtttttaaaaaataggataaaatttttttctataaaaaataaa
aaacatggaaaaatagggaaatgctgtttttgaggtaatttaataatacacagaatttttcattagtgtcggaaggtctaaaaagacaaaagtatatcatgg
gaataaaaaaagatagaaaaggaaaacagtttaggaatttgcttaacaaatgaaaaatgccttttaaaatggcatcagtcgaagcaagtgtgtgtgcatt
tattatatgtccaaaataaaatgctaatttcataaaaattaagg

Figure 7B

cgtggcgacgcgtgtagtagtccagctacttggaggctgaggcagaagaatcccttgaacccaggagcgagggttgacagtaagccaagattgtc
 actgcactccagcctggcgaaagaggaagactccatctcaaaaaaagaaaaatctcatggttatgcaactcttattatgatcagaaaaat
 ggacattttagtattaaactctgtacaatgtttcatgtagtaaaaaataataaaaaactattaatcatctagcttggagagataggagaaagacatta
 ctgtcactagtcaaatatataatctttactatccacaaaaatctctctgattctggttagaaggcatactattaaatgataaaaaataaaact
 gaaggcctcaacatatcacagggtataaagaatatagggaagttagttcaatagtttaaaataaagcacactcttacagtataagaactagtcggg
 cttttatgcctttagtttagttcttactcttcccttaactcttttctgttgatgtaattacattaatgcttaagagtgaactttttaagtgtgggta
 aaaaagaaataattacttacaaaagtttaattcttccatttcccttgagagaggaagttatgaaaagcagctcttatctaaagcaaaagagccacag
 attgatttcatagggtgacagaaactacccagttgtgcatgtctgatgttaattcacataatgaATGTATGAATTAATGTTCTTATTCATGTTGATACA
 GCCTCAGTCCATGGCGCATCCGTGTGGGGACCCCAACATACCCAGAAATCACAGATATTTTCCCAACTATTCATGAACGTCCAGTTCTTTTTCAC
 CACCTCCACCTGCCACCGAAAGTAGCCATTTCCAGCGGGCTAAGAGCACCTCTCTCTGGAAGCCCCAACTCACCACTTCCAACCCCTGCTGAGG
 ACTGTTGGCCAAAGTCTTCTCCACCTGGTGGCAGCCCAACTAACTGACACCCAGAGCCGTAGTTATGTTGGGTACTACAGCCAGTAGAGTAACCTGG
 AGAGTCATGTGAGATACAGGTCCATCCTATGTTTGAACCATCTCAAGTTTACAGTGACTATAGACCTGGACTAGTACTTCCAGAAGAACTCACTATT
 TTATTCCTCAGGAAGCAGTGTATGTAGCTGGGTACATTACAGGCCCGGGTGGCAGAACAGTATGAGGGCATTCATACAACTCATCAGTACTGTCA
 AGTCCATGAACAGATACCTGAACAGAAGCCAGTACAAAGGGGCCCTACTTCAAGTTCTGTCTTTGAATTCATCTGGACAGGCTTTCCTGGTAGG
 ACACCTTCAGAAATCTAAGATTAGATTCTGGATTGGTCCGGGATCTCCCTCTCTAGTATTTCTGCACCTATCAGTACAGATGCTACACGTTTGAAAT
 TTCACCTGTCTTTGTTCCCTCATTTCTGCGCCCTGCTGTGTTAACTCATAAACAATGAGAGCAGAACTGTGTATTTGAATTTTCATGTTTCACACACCA
 AGTCCCTCTTCAGGAGAAGGAGGTGGAATTTTACCTCAGCGTGTACTGCGAAATCGGCAGGTTCAGTGGACTTGAATCAAGAAGAACTGCCTCCTCA
 AATCAGTTGGATTACATGGCTACTTGCAGCCTGTGACTGAAGAAAGCATAATTACCATGCCCCAGAAATGACCGTTTCTGTGGTAGAGCCTATCGGA
 CAGAACTGGCCAAATAGGAAGCCAGAAATATCCAGTGATTCTTCACAAATCACTTCTTCAGACCCAGTGATTTTCAGTCACTCCCCCTACAGGGGG
 AGCAGCTGCACCTTTTGGCTCTGACGTCTCAATGCCCCTTATCCATCTGCCCTCAGACAGTGTACAAAGAAATCCCCACTTTTCTGTGTTTCCCCCAAG
 GAACCACTCAGCAGGTCTTAACTGCCTCATTTTCTTCAGGAGGATCTGCACCTTCATCCACAGGTTATAGGAAAACTTCCACAAATATTTTAAact
 accctactttgcaccataaacatttaaatttctatttccctgaatcatggattttggagaaatattgtttaattttatcagtagagtctcc
 ccacttttgggggttgtgaactacatatatgcatgttaaaaaaataatgtgagaaagctacctgatttaccattattatgtgaaaccagtggaataa
 aacacaaaaactagaatttttagtcattcttcacaatcacgacttctatgcacgttattttcaaccagtagtgaataatgcaatgtgaaatgtaataatg
 ttgacccagcattatttaggaatacaaatcttaagtattacttcttcccaacaagtttttaaaaaataggataaaattttttctataaaatat
 aaaaatggaataatggaatgctgtttttgagggttaataataacacagaatttttcatttagtgcgaaggatctaaaaagacaaaagtatatcat
 ggggaataaaaaagatagaaaaaggaacagtttaggaatttgccttaacaaaatgaaaatgccttttttaaaatggcatcagccaagcaagtgtgctgtgc
 attattatatgtccaaaataaaatgctaatttcataaaattaagg

Figure 7C

cgtggtggcgacgcgtgtagtcccagctacttgggaggtgagggcagaagaatcccttgaacccagggaggcgggaggttgcagtaagccaagattgtc
 actgcactccagcctggcgaaagaggaagactccatctcaaaaaaagaaaaaaatattcatggttatgcaactcttatttatgatcagaaaaat
 ggacattttagtatttaactctgaacatggttcatgtagtaaaaaataataaaactattaatcatctagcttgggagagataggagaaaaagacatta
 ctgtcactagtcacaaattatataatctttactatccacccaaaaatctctctgatttctggttagaaggcatactatttaattgataaagaaaaataaaact
 gaaggccttaacatatcacagggtaataagaatatagggaaagttagttcaatagtttaaaataaagcacacacttctacagtataagaactagtcggg
 ctttatgccttgtttagttcttactcttcccttaactctttctgttgatgtaattacattaatgcttaagagtgaactttttaagtgtgggtg
 aaaaagaaataaattacttacaagtttaattcttccatttcccttgagagaggaagttaaggaaaagcagctcttataaagcaaaagagccacag
 attgatttccattggcctggatgtatttaattgatttttactatgcacataaattccagaagcattgttatttattataataatttaatttagtga
 accatttcataagggttacacagaactaccagttgtgcatgtctgatgtaatttcacatatgaATGTATGAATTACTTGTCTTATTTCATGTTGATACA
 GCCTCAGTCCATGGCGCATCCGTGTGGGGACCCCAACATACCCAGAAATCACAGATATTTTCCCAACTATTTCATGAACGTCAGTTCTTTTTCAC
 CACCTCCACCTGCCACCCGAAAGTAGCCATTTCACAGCGGGTAAGAGCACCTCTTCTGGAGCCCAAACTCACCACTTCCAACCCCTGCTGAGG
 ACTGTTGGCCAAAGTCTTCTTCCACCTGGTGGCAGCCCAACTAAGTGGACACAGAGCCGTAGTTATGTTGGGTACTACAGCCAGTAGAGTAACCTGG
 AGAGTCATGTGAGATACAGGTCATCCTATGTTTGAACCATCTCAAGTTTACAGTGACTATAGACCTGGACTAGTACTTCCAGAAGAAAGCTCACTATT
 TTATTCTCAGGAAGCAGTGTATGTAGTGGGTACATTACCAGGCCGGGTGGCAGAACAGTATGAGGGCATTCATACAACTCATCAGTACTGTCA
 AGTCCATGAACACAGATACCTGAACAGAACCCAGTACAAAGGGGCCCTACTTCAAGTTCTGTCTTTGAATTTCCATCTGGACAGGCTTCCCTGGTAGG
 ACACCTTCAGAACTAAGATTAGATTCTGGATTGGTCCGGGATCTCCCTCTCTAGTATTCTGCACCTATCAGTACAGATGCTACACGTTTGAAT
 TTCACCTGTCTTGTTCCTCATCTCGCCCTGCTGTGTAACTCATAAACATGAGAGCAGAAGCAACTGTGTATTGAATTTTCATGTTTCACACACCA
 AGCTCCTCTTCAGGAGAAGGAGGTGGAATTTTACCTCAGCGTGTTTACCGAAATCGGCAGGTGTCAGTGGACTTGAATCAAGAAGAACTGCCTCCTCA
 ATCAGTTGGATTACATGGCTACTTGTAGCCTGTGACTGAAGAAAAGCATAATTACCATGCCCCAGAAATTGACCCGTTTCTGTGGTAGAGCCTATCGGAC
 AGAACTGGCCCAATAGGAAGCCAGAAATATCCAGTGATTCCCTCAAAATCACTTCTTCAGACCCAGTGATTTTCAGTCACCTCCCCCTACAGGGGA
 GCAGCTGCACCTTTTGGCTCTGACGTCTCAATGCCCTTTATCCATCTGCCCTCAGACAGTGTACAGAATCCCCACTTTTCTTCTGTTTCCCCCAAGG
 AACCACATCTCAGCAGGTCTTAACCTGCCCTCATTTTCTTCAGGAGGATCTGCACCTTCATCCACAGGTTATAGGAAAACCTCCACAATTATTTTAAacta
 ccctacttggcaccataaacatttaaaatttctatcccttatttccctgaatcatggatttggagaaaatttgaatttattatcagtagagatttccc
 catcttgggggttgaactacatatatgcatttcaaaaaaataatgtgagagaagctacctgatttaccattatattgtgaaaaaccagtggaataa
 acacaaaaactagaattttagtcattcttcacaaatcacgacttctatgcagttattttcaaccagtagtgaaaatgcaagtgatgtaatgtatggt
 tgaccagcattatttaggaatacaaaatcttaagtattacttctcctccaaaagtttttaaaaaataggataaaatttttttctataaaaaata
 aaacatggaataatagggaatgctgtttttgaggttaataataacacagaatttttcatttagtgcgaaggatctaaaaagacaaaagtatatcatg
 ggaataaaaaagatagaaaaaggaaaacagtttaggaatttgccttaacaaatgaaaatgccttttttaaaatggcatcagtaagcaagttgctgtgca
 ttattatatgtccaaaataaaatgctaatttcataaaaaattaagg

nt+46



Figure 9

Alignment with human hsn2pep:

```

      *           20           *           40           *
pESTframe+ : MYELLVLFMLIQPQSMABRGGTPTYESQIFFPTIHEREVVSFSPEPTCF : 50
human       : MYELLVLFMLIQPQSMABRGGTPTYESQIFFPTIHEREVVSFSPEPTCF : 50
              MYELLVLFMLIQPQSMABRGGTPTYESQIFFPTIHEREVVSFSPEPTCF

      60           *           80           *           100
pESTframe+ : PKVAISQRRKSTSFLAQTHHFQPLLRTVGQQLLPFGGPTNWTPEAVVM : 100
human       : PKVAISQRRKSTSFLAQTHHFQPLLRTVGQQLLPFGGPTNWTPEAVVM : 100
              PKVAISQRRKSTSFLAQTHHFQPLLRTVGQQLLPFGGPTNWTPEAVVM

      *           120           *           140           *
pESTframe+ : LGTTASRVGTGECEIQVQPLFEPTQVYDPRGLVLPPEAAHYFIPQEA : 150
human       : LGTTASRVGTGECEIQVQPLFEPTQVYDPRGLVLPPEAAHYFIPQEA : 150
              LGTTASRVGTGECEIQVQPLFEPTQVYDPRGLVLPPEAAHYFIPQEA

      160           *           180           *           200
pESTframe+ : VAGVHYQTMAEQFEGIPYNSVLSSPMKQIPEQKPVQGGPSSSVFEFF : 200
human       : VAGVHYQTMAEQFEGIPYNSVLSSPMKQIPEQKPVQGGPSSSVFEFF : 200
              VAGVHYQTMAEQFEGIPYNSVLSSPMKQIPEQKPVQGGPSSSVFEFF

      *           220           *           240           *
pESTframe+ : SGQAFLVGHLQNLRLDSGLPGSPLSSISTPISTDATRLKFHEVFPVPHSA : 250
human       : SGQAFLVGHLQNLRLDSGLPGSPLSSISTPISTDATRLKFHEVFPVPHSA : 250
              SGQAFLVGHLQNLRLDSGLPGSPLSSISTPISTDATRLKFHEVFPVPHSA

      260           *           280           *           300
pESTframe+ : PAVLTHNNESRSNCVFEEHVHTPSSSSGEGGVLPQRIYRNRQVAVDLNQ : 299
human       : PAVLTHNNESRSNCVFEEHVHTPSSSSGEGGVLPQRIYRNRQVAVDLNQ : 300
              PAVLTHNNESRSNCVFEEHVHTPSSSSGEGGVLPQRIYRNRQVAVDLNQ

      *           320           *           340           *
pESTframe+ : EEPFPPQSGLHGGLQPVTEEQHNFQPELTVSVVEPTGQSMPIGSPPEYSS : 349
human       : EEPFPPQSGLHGGLQPVTEEQHNFQPELTVSVVEPTGQSMPIGSPPEYSS : 350
              EEPFPPQSGLHGGLQPVTEEQHNFQPELTVSVVEPTGQSMPIGSPPEYSS

      360           *           380           *           400
pESTframe+ : DSSQITSSDPSDFQSPPEPTGCTAAFFGSDVSLPFIHLPQTVIQESPLFFC : 399
human       : DSSQITSSDPSDFQSPPEPTGCTAAFFGSDVSLPFIHLPQTVIQESPLFFC : 400
              DSSQITSSDPSDFQSPPEPTGCTAAFFGSDVSLPFIHLPQTVIQESPLFFC

      *           420           *
pESTframe+ : FPQGTTSQILSASFSSGGSSALHPQVIGKLPQS : 433
human       : FPQGTTSQVLTASFSSGGSSALHPQVIGKLPQS : 434
              FPQGTTSQVLTASFSSGGSSALHPQVIGKLPQS

```


SEQUENCE LISTING

<110> Xenon Genetics Inc.

<120> HSAN II Related Gene and Expression Products
and Uses Thereof

<130> 760050-110

<140>

<141>

<150> US/60/425,601

<151> 2002-11-12

<150> US/60/502,453

<151> 2003-09-12

<160> 33

<170> PatentIn version 3.0

<210> 1

<211> 2786

<212> DNA

<213> Homo sapiens

<400> 1

cgtgggtggcg	cacgcgtgta	gtcccagcta	cttggggaggc	tgaggcagaa	gaatcccttg	60
aacccaggag	gcggagggtt	cagtaagcca	agattgtcac	tgcactccag	cctgggcgaa	120
agaggaagac	tccatctcaa	aaaaaagaaa	aaagaaat	catgggttat	caactcttat	180
ttatgatcag	aaaaatggac	atgttgtgat	ttaactctgt	aacatgtttc	atgtagtaaa	240
aatataataa	aactattaat	catctagctt	gggagagata	ggagaaaagac	attactgtca	300
ctagtcaaat	tatatatctt	ttactatcca	ccaaaaatct	cttctgattt	ctgggttagaa	360
ggcatactat	taattgataa	gaaaataaaa	ctgaaggcct	ctaacataatc	acagggtaat	420
aagaatatag	ggaaagttag	ttcaatagtt	ttaaattaaag	cacacttctt	acagtataga	480
actagtggg	cttttatgcc	ttgttttagt	tcttactctt	cctttaactc	tttttctgtt	540
gatgtaattt	acattaatgc	ttaagagtga	actttttaag	tgtgggtaaa	aacgaaataa	600
ttacttacaa	agtttaattc	ttccatttcc	tttgagagag	gaaagttatg	gaaaagcagc	660
tcttatctaa	agcaaagagc	ccacagattg	atttcattgg	ccctggatgt	atttaattgga	720
tttttactat	gcacataatt	tccagaagca	ttgttattta	tttatttaatt	ataaatttag	780
tgtaacatt	tcatagggtt	acacagaact	accagttgt	gcatgtctga	tgtaatctca	840
catatgaatg	tatgaattac	ttgtcttatt	catgttgata	cagcctcagt	ccatggcgca	900
tccgtgtggg	gggaccccaa	catacccaga	atcacagata	tttttcccaa	ctattcatga	960
acgtccagtt	tctttttcac	cacctccac	ctgcccaccg	aaagtagcca	tttcccagcg	1020
gcgtaagagc	acctccttcc	tggaaagcca	aactcaccac	ttccaacccc	tgctgaggac	1080
tggtggccaa	agtcttcttc	cacctggtgg	cagcccaact	aactggacac	cagaggccgt	1140
agttatgttg	ggtactacag	ccagtagagt	aactggagag	tcatgtgaga	tacaggtcca	1200
tcctatgttt	gaaccatctc	aagtttacag	tgactataga	cctggactag	tacttccaga	1260
agaagctcac	tattttattc	ctcaggaagc	agtgtatgta	gctggggtac	attaccaggc	1320
ccgggtggca	gaacagtatg	agggcatttc	atacaactca	tcagtactgt	caagtcctat	1380
gaaacagata	cctgaacaga	agccagtaca	agggggccct	acttcaagtt	ctgtctttga	1440
atttccatct	ggacaggctt	tcctggtagg	acaccttcag	aatctaagat	tagattctgg	1500
attgggtccg	ggatctcccc	tctctagtat	ttctgcacct	atcagtacag	atgctacacg	1560
tttgaaattt	cacctgtct	ttgttcctca	ttctgcgcct	gctgtgttaa	ctcataacaa	1620
tgagagcaga	agcaactgtg	tatttgaatt	tcatgttcac	acaccaagct	cctcttcagg	1680
agaaggaggt	ggaattttac	ctcagcgtgt	ttaccgaaat	cggcagggtg	cagtggactt	1740
gaatcaagaa	gaatgcctc	ctcaatcagt	tggattacat	ggctacttgc	agcctgtgac	1800
tgaagaaaag	cataattacc	atgcccaga	attgaccgtt	tctgtggtag	agcctatcgg	1860
acagaactgg	ccaataggaa	gccagaata	ttccagtgt	tcctcacaaa	tcacttcttc	1920

```

agaccccagt gattttcag cacctcccc tacaggggga gcagctgcac ctctggctc 1980
tgacgtctca atgcccttta tccatctgcc tcagacagtg ttacaagaat ccccactttt 2040
cttctgtttc ccccaaggaa ccacatctca gcaggtctta actgcctcat tttcttcagg 2100
aggatctgca cttcatccac aggttatagg aaaacttcca caattatttt aaactaccot 2160
actttgcacc ataacattta aattttctat tccttatttc cctgaatcat ggattttgga 2220
gaaatattgt ttaattttat cagtagagtt tccccatctt tgggggggttg tgaactacat 2280
atatgcattt aaaaacaaaa tgtgagagaa gctacctgat ttacctatta tatgtgaaaa 2340
ccagtggaaa aaacacaaaa actagaattt tagtcattct tcacaatcac gacttctatg 2400
cacgttattt tcaaccagta gtgaaaatgc aagtgtatgt aatgtatggt tgacccagca 2460
ttatttagga atacaaatct taagtattac tttcttcctc caaacaagtt tttaaaaaat 2520
aggataaatt ttttttctat aaaatataaa acatggaaaa tagggaatgc tgtttttgag 2580
gtaatatata taatacacag aattttcatt agtgtcgaag gatctaaaaa gacaaagtat 2640
atcatgggaa taaaaaaaga tagaaaagga aacagtttag gaatttgcct taacaaatga 2700
aaatgccttt ttaaaatggc atcagtcgaag caagttgctg tgcattatta tatgtccaaa 2760
taaaatgcta attcataaaa ttaagg 2786

```

<210> 2
 <211> 434
 <212> PRT
 <213> Homo sapiens

<400> 2

```

Met Tyr Glu Leu Leu Val Leu Phe Met Leu Ile Gln Pro Gln Ser Met
1          5          10          15

Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile Phe
20          25          30

Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Pro Thr
35          40          45

Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser Phe
50          55          60

Leu Glu Ala Gln Thr His His Phe Gln Pro Leu Leu Arg Thr Val Gly
65          70          75          80

Gln Ser Leu Leu Pro Pro Gly Gly Ser Pro Thr Asn Trp Thr Pro Glu
85          90          95

Ala Val Val Met Leu Gly Thr Thr Ala Ser Arg Val Thr Gly Glu Ser
100         105         110

Cys Glu Ile Gln Val His Pro Met Phe Glu Pro Ser Gln Val Tyr Ser
115         120         125

Asp Tyr Arg Pro Gly Leu Val Leu Pro Glu Glu Ala His Tyr Phe Ile
130         135         140

Pro Gln Glu Ala Val Tyr Val Ala Gly Val His Tyr Gln Ala Arg Val
145         150         155         160

Ala Glu Gln Tyr Glu Gly Ile Pro Tyr Asn Ser Ser Val Leu Ser Ser
165         170         175

Pro Met Lys Gln Ile Pro Glu Gln Lys Pro Val Gln Gly Gly Pro Thr
180         185         190

Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val Gly
195         200         205

```

His Leu Gln Asn Leu Arg Leu Asp Ser Gly Leu Gly Pro Gly Ser Pro
 210 215 220
 Leu Ser Ser Ile Ser Ala Pro Ile Ser Thr Asp Ala Thr Arg Leu Lys
 225 230 235 240
 Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr His
 245 250 255
 Asn Asn Glu Ser Arg Ser Asn Cys Val Phe Glu Phe His Val His Thr
 260 265 270
 Pro Ser Ser Ser Ser Gly Glu Gly Gly Gly Ile Leu Pro Gln Arg Val
 275 280 285
 Tyr Arg Asn Arg Gln Val Ala Val Asp Leu Asn Gln Glu Glu Leu Pro
 290 295 300
 Pro Gln Ser Val Gly Leu His Gly Tyr Leu Gln Pro Val Thr Glu Glu
 305 310 315 320
 Lys His Asn Tyr His Ala Pro Glu Leu Thr Val Ser Val Val Glu Pro
 325 330 335
 Ile Gly Gln Asn Trp Pro Ile Gly Ser Pro Glu Tyr Ser Ser Asp Ser
 340 345 350
 Ser Gln Ile Thr Ser Ser Asp Pro Ser Asp Phe Gln Ser Pro Pro Pro
 355 360 365
 Thr Gly Gly Ala Ala Ala Pro Phe Gly Ser Asp Val Ser Met Pro Phe
 370 375 380
 Ile His Leu Pro Gln Thr Val Leu Gln Glu Ser Pro Leu Phe Phe Cys
 385 390 395 400
 Phe Pro Gln Gly Thr Thr Ser Gln Gln Val Leu Thr Ala Ser Phe Ser
 405 410 415
 Ser Gly Gly Ser Ala Leu His Pro Gln Val Ile Gly Lys Leu Pro Gln
 420 425 430
 Leu Phe

<210> 3
 <211> 206
 <212> PRT
 <213> Homo sapiens

<400> 3
 Met Tyr Glu Leu Leu Val Leu Phe Met Leu Ile Gln Pro Gln Ser Met
 1 5 10 15
 Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile Phe
 20 25 30
 Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Pro Thr
 35 40 45

Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Phe
 50 55 60
 Leu Glu Ala Gln Thr His His Phe Gln Pro Leu Leu Arg Thr Val Gly
 65 70 75 80
 Gln Ser Leu Leu Pro Pro Gly Gly Ser Pro Thr Asn Trp Thr Pro Glu
 85 90 95
 Ala Val Val Met Leu Gly Thr Thr Ala Ser Arg Val Thr Gly Glu Ser
 100 105 110
 Cys Glu Ile Gln Val His Pro Met Phe Glu Pro Ser Gln Val Tyr Ser
 115 120 125
 Asp Tyr Arg Pro Gly Leu Val Leu Pro Glu Glu Ala His Tyr Phe Ile
 130 135 140
 Pro Gln Glu Ala Val Tyr Val Ala Gly Val His Tyr Gln Ala Arg Val
 145 150 155 160
 Ala Glu Gln Tyr Glu Gly Ile Pro Tyr Asn Ser Ser Val Leu Ser Ser
 165 170 175
 Pro Met Lys Gln Ile Pro Glu Gln Lys Pro Val Gln Gly Gly Pro Thr
 180 185 190
 Ser Ser Ser Val Phe Asp Phe His Leu Asp Arg Leu Ser Trp
 195 200 205

<210> 4
 <211> 318
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Tyr Glu Leu Leu Val Leu Phe Met Leu Ile Gln Pro Gln Ser Met
 1 5 10 15
 Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile Phe
 20 25 30
 Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Pro Thr
 35 40 45
 Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser Phe
 50 55 60
 Leu Glu Ala Gln Thr His His Phe Gln Pro Leu Leu Arg Thr Val Gly
 65 70 75 80
 Gln Ser Leu Leu Pro Pro Gly Gly Ser Pro Thr Asn Trp Thr Pro Glu
 85 90 95
 Ala Val Val Met Leu Gly Thr Thr Ala Ser Arg Val Thr Gly Glu Ser
 100 105 110
 Cys Glu Ile Gln Val His Pro Met Phe Glu Pro Ser Gln Val Tyr Ser
 115 120 125

Asp Tyr Arg Pro Gly Val Leu Pro Glu Glu Ala His Tyr Ile
 130 135 140
 Pro Gln Glu Ala Val Tyr Val Ala Gly Val His Tyr Gln Ala Arg Val
 145 150 155 160
 Ala Glu Gln Tyr Glu Gly Ile Pro Tyr Asn Ser Ser Val Leu Ser Ser
 165 170 175
 Pro Met Lys Gln Ile Pro Glu Gln Lys Pro Val Gln Gly Gly Pro Thr
 180 185 190
 Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val Gly
 195 200 205
 His Leu Gln Asn Leu Arg Leu Asp Ser Gly Leu Gly Pro Gly Ser Pro
 210 215 220
 Leu Ser Ser Ile Ser Ala Pro Ile Ser Thr Asp Ala Thr Arg Leu Lys
 225 230 235 240
 Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr His
 245 250 255
 Asn Asn Glu Ser Arg Ser Asn Cys Val Phe Glu Phe His Val His Thr
 260 265 270
 Pro Ser Ser Ser Ser Gly Glu Gly Gly Gly Ile Leu Pro Gln Arg Val
 275 280 285
 Tyr Arg Asn Arg Gln Val Ala Val Asp Leu Asn Gln Glu Glu Leu Pro
 290 295 300
 Pro Gln Ile Ser Trp Ile Thr Trp Leu Leu Ala Ala Cys Asp
 305 310 315

<210> 5
 <211> 314
 <212> PRT
 <213> Homo sapiens

<400> 5
 Met Tyr Glu Leu Leu Val Leu Phe Met Leu Ile Gln Pro Gln Ser Met
 1 5 10 15
 Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile Phe
 20 25 30
 Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Thr
 35 40 45
 Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser Phe
 50 55 60
 Leu Glu Ala Gln Thr His His Phe Gln Pro Leu Leu Arg Thr Val Gly
 65 70 75 80
 Gln Ser Leu Leu Pro Pro Gly Gly Ser Pro Thr Asn Trp Thr Pro Glu
 85 90 95

Ala Val Val Met Leu Tyr Thr Thr Ala Ser Arg Val Thr Gly Ser
100 105 110

Cys Glu Ile Gln Val His Pro Met Phe Glu Pro Ser Gln Val Tyr Ser
115 120 125

Asp Tyr Arg Pro Gly Leu Val Leu Pro Glu Glu Ala His Tyr Phe Ile
130 135 140

Pro Gln Glu Ala Val Tyr Val Ala Gly Val His Tyr Gln Ala Arg Val
145 150 155 160

Ala Glu Gln Tyr Glu Gly Ile Pro Tyr Asn Ser Ser Val Leu Ser Ser
165 170 175

Pro Met Lys Gln Ile Pro Glu Gln Lys Pro Val Gln Gly Gly Pro Thr
180 185 190

Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val Gly
195 200 205

His Leu Gln Asn Leu Arg Leu Asp Ser Gly Leu Gly Pro Gly Ser Pro
210 215 220

Leu Ser Ser Ile Ser Ala Pro Ile Ser Thr Asp Ala Thr Arg Leu Lys
225 230 235 240

Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr His
245 250 255

Asn Asn Glu Ser Arg Ser Asn Cys Val Phe Glu Phe His Val His Thr
260 265 270

Pro Ser Ser Ser Ser Gly Glu Gly Gly Gly Ile Leu Pro Gln Arg Val
275 280 285

Tyr Arg Asn Arg Gln Val Ala Val Asp Leu Asn Gln Glu Glu Leu Pro
290 295 300

Pro Gln Ser Val Gly Leu His Gly Tyr Leu
305 310

<210> 6
<211> 5562
<212> DNA
<213> Homo sapiens

<400> 6
gtaagtaa at gcttaagagc atgtaatact acaatgagag ccaatggata gtctgctaaa 60
ttaaaattct ttgtacaaac caagtaacag tatgagtctg aagtagaaaa tataaaatgc 120
tgtgtctcta aagccttatt atagagtagg atatagacat taatgtgtgg gcattagctg 180
aactgcctac atgaagcagc tatgtgcaag atttagatta gtctgcattt ttactaaagg 240
gaacgctctc ttttagcaactt aagatgttga aactattcac agcaacaaaa gctggacata 300
caaagtgttt gcggtgtggg cgagaagagt ttttcaaaag cttgttaaaa tgattactaa 360
gtatttcctg taagattccc tattgtacaa aacacaggta agaccttcaa accattatag 420
aatgctgcac gtgagtactc tagattgagg gaatccttta tttgaatgac tgtttaattt 480
agcgttgtg ttgagttaat catcagttgc tgtgtgtgtt cctgcctttt aaagtctgtg 540
ttggacacaa gttccttccc cagtatatgt tcttttgctt tttcccttta gtttgtaact 600
gttttattac tgatttattt tcagaatcaa cttatagaaa ttatttcctg aaacatttta 660
tgcttgtata cacatacaca cacacaatta tcctgtgttt tacattagac actcttgttt 720

ctcatataga	gaatgtt	gtagatcatc	agataatacc	ccatttcttt	tgttgacctc	780
agaggaaaat	aaaatacgtc	atctgcctct	gaaaggtatt	agaaaataag	taacttgtca	840
ctgtacagca	cttgtttagt	gagatcttta	agcagacaga	ttttccctt	ctcagaaagc	900
ttttaatcta	aactttgttg	aattccttgct	cttctttaaa	tgaaaatctt	tctctgttaa	960
gaaaataactt	ctgcttggtc	gatttttttat	cctacagctt	ggaataaaaag	atgttcattt	1020
gtttcagagt	cogtttaatc	ctgctttggc	taggtctcaa	aatgaatggc	cacagtaaca	1080
caaataaaaa	actaatatac	taacttttta	aaacaatcta	ctttctgtct	ttaaaataat	1140
tgagggtaga	atagaagaac	tttataaaga	ctgcacagaa	aatataatttc	tgtgttctga	1200
caaataattac	ttgtcagcat	ttcttgtcag	ttctacctt	taattgtttt	tttctggat	1260
tattataact	agattttgct	aagacttttt	ttttaaaatt	ttctcttgct	tatactcatc	1320
atgggggagc	attaggtctc	tatttttagaa	agcaagtaga	atgacatgaa	gatattattt	1380
gagccattaa	tctagctata	tttattgatg	ctctaaaagc	agacgttctt	gctttttaat	1440
gacagtattc	tatggaggta	ttatgaagaa	ttattaaata	ttgagaattt	aaaaatttgt	1500
ctaattctat	tcttgtctca	tttcaggaat	gtggcttaca	ataatatgga	caaattataa	1560
gtgaaatttt	agaaattaga	ataaaaagaa	taaggctaga	gagaaatcaa	aacaggaatg	1620
aaactaaaca	atcatatgat	gatatgtaaa	tgttggtat	tggtgggcca	cagatttggc	1680
ccaaagcatt	ctacacagcc	aagttttgga	aatgaaaaca	tagtttcata	gttcacaaca	1740
tctaaaagga	aaataacagt	cttattcaag	acaaccttta	tgggtattgat	catactgcct	1800
tatagccata	ttaaacctg	tgcctgaaca	tttttttaag	tttttttttt	tttttttgac	1860
agcgtctcac	tctgtctccc	aggtgggaat	acagtggcgt	gatcttggct	cactgcaact	1920
ttcacctcct	gggttcaagt	gatttctcctg	cctcagcctc	cccagtagct	aagattacag	1980
gcacccatca	ccacgcccag	ctaatttttg	tatttttttt	agtagagaca	cagtttcacc	2040
tgtgttgcta	ggctggctct	gaactcttga	cctcaagtga	tccacctgcc	tccgacctccc	2100
aaagtgccag	gattacaggt	gtgagccacc	atgcccgcc	taaattatta	tttagagatt	2160
cagattgaag	atattctccc	attgattact	ttgtaaaagt	acaaaatatt	tcctatatcc	2220
agttgtgaat	ttcttctaga	aagctaaact	ttttcctagt	agatcatgac	ttttttacct	2280
ttacaaataa	agagtggcaa	gtataagaaa	tactgtgagc	caacaacttt	ccatctcatg	2340
tctttataaaa	tagaaatgag	aatctttttc	tgagggttcc	agagctcttg	caaaataatc	2400
caatcttaaa	attctctgta	tcttttaattg	tactttaaac	taatagaatt	cattcttgtt	2460
gaatcttaag	tttgccaaag	ttgcaagcta	atattcattg	ttttccctgt	ctgcacctt	2520
cagtcagttg	gtcattctta	aattcataca	cttgtaacac	ttgggtgttat	ttcagaatat	2580
ttagcatgat	aactgatcat	tatcatgata	attcattaaa	cctgtccttg	actgacaacg	2640
cagaatagtt	gtgaatagtt	gtgatctgtc	agatgtgcta	tgttgagtta	tagtctcct	2700
caaaaaaaaa	actgactttg	tgggaattggg	agaggatgga	acatttcttc	atgcaactat	2760
gccctgttat	cattgtactt	ttgtttattt	tgttttcaca	gtactgtgtt	tttcatgtgt	2820
gtgtttgttt	tgtgttgagc	ctcgtcgtgg	cogtagcatg	tcgggttgtg	ttcccatctt	2880
tctgctgttg	cctctgtgtc	ccgcatctct	cccagtgtc	ttccacccca	ccgccagtac	2940
tgtctgcacc	tctttctcct	tcctcctccc	ggactgcccc	gaggaaactt	ttgccgaaaa	3000
gctttctaaa	gcattggaga	gtgtcctgcc	tactgactct	gcctctcagc	gcaagcaccg	3060
acgctccagc	ctgccttccc	tctttgtcag	tactgtatgt	aactgtaaac	ttctgacaaa	3120
tgaacaatta	ttaccaatga	atacatccag	gcaacaagaa	ttttaattaa	aattgaataa	3180
agaacaggac	taaacttggt	tatattatgc	ttctaggata	ccagttattc	ctgtaggaaa	3240
cttttttgta	ataaggcttt	aagggggaat	aaaaagcaag	gaggtaatga	gagagaatac	3300
ctataatgtc	cttggcacat	accagcatgt	gtccaataaa	atttaccttt	cattatatct	3360
tctgttttcc	tatcgacata	aaattctttc	attttcatat	tctagtttta	aaaggattaa	3420
ttttaaagtt	attgcottaa	tcaaaagcaa	attttatcaa	acattaccca	aacatttttt	3480
ctcctcaact	tattctaagg	agacactaat	aaaactggat	gagtcctttt	tttttttaatt	3540
aaccttagtc	agcaaagaat	aaccttggag	tgaacattta	ttttaattta	atcatttatg	3600
taagatgatg	tatcagcaaa	ccaaagaact	acataatgtt	cccccatgga	agtatttttt	3660
ccttaggttt	caacctaaag	tgggactgta	gctctaata	atgcattaga	tgcataatata	3720
gtatgtatga	ctttcttttg	ctccattaat	gtttataaaa	gatctatctt	ataaaaattat	3780
tttaagcatg	tgggtccttg	ctgcttttcc	aaataaagaa	tatagttaac	tttctgaact	3840
agcaacccca	gaattttcag	atttttactt	aggagaatga	agaactatac	agtgttatta	3900
tggctttgag	acataccatg	agtaaaatac	aataaaaatga	gcattctgac	tgaaaaataga	3960
aatccttatt	tagctattgt	gaacagttga	tctactttgc	cagctcgata	aggatcttcc	4020
ctatggaaga	atgaaactga	attcttttaa	aaagcggggg	aggctgtgat	agtgaattct	4080
tactattatt	tacttattta	tttatgtttt	ggagatggaa	tctcgctctg	tcgccaggct	4140
ggagtgcagt	ggcacgatct	cgactcacgg	caacctctgc	ctcctgggtt	caagtgattc	4200
tctgcctca	gcctcccgag	tagctaggac	tacaggcgca	tgccaccaca	cccagctaata	4260
ttttgtattt	tagtagagat	gggtttcacc	atgtttggcc	aggatggtct	tgatctcctg	4320
acattgtgat	ccacccctt	cagcctccca	cagtgtcggg	attacaggcg	tgagccaccg	4380

cgcccagcgg	gtaatgaa	attatagtg	atattgtttg	catttgccaa	gcgtgtttc	4440
taaatttagt	cttaaatcta	gaactgcttc	tagaataaaa	agtgcctatca	aataaaattg	4500
gctgtcattt	atattaccct	tttcagttga	ttagaacacc	aggatgtcag	aatcagaaaa	4560
ttcatgggta	cggccgggct	cagtggctca	cgctgtaat	cctagcactt	tgggaggccg	4620
aagtgaagcag	atcacctgag	gtcaggagtt	cgagaccagc	ctggccaaca	tggcgaaatg	4680
tgtctctact	aaaaatacaa	aaattagctg	ggcgtgggtg	cgacgcgtg	tagtcccagc	4740
tacttgggag	gctgaggcag	aagaatccct	tgaacccagg	aggcggaggt	tgcagtaagc	4800
caagattgtc	actgcaactcc	agcctgggcg	aaagaggaag	actccatctc	aaaaaaaaga	4860
aaaaagaaat	ttcatgggta	tgcactctt	atattatgatc	agaaaaatgg	acattttgtg	4920
atttaactct	gtaacatggt	tcatgtagta	aaaatataat	aaaactatta	atcatctagc	4980
ttgggagaga	taggagaaaag	acattactgt	cactagtcaa	attatatatc	ttttactatc	5040
cacaaaaaat	ctcttctgat	ttctgggttag	aaggcatact	attaattgat	aagaaaaataa	5100
aactgaaggc	ctctaacata	tcacagggta	ataagaatat	agggaaagtt	agttcaatag	5160
tttaaattaa	agcacacttc	ttacagtata	gaactagtgc	ggcttttatg	ccttgtttta	5220
gttcttactc	ttcctttaac	tctttttctg	ttgatgtaat	ttacattaat	gcttaagagt	5280
gaacttttta	agtgtgggta	aaaacgaaat	aattacttac	aaagttaaat	tcttccattt	5340
cctttgagag	aggaaagtta	tggaaaagca	gctcttatct	aaagcaaaga	gcccacagat	5400
tgatttcatt	ggccctggat	gtatttaaatg	gatttttact	atgcacataa	tttccagaag	5460
cattgttatt	tatttattaa	ttataaattt	agtgtaacca	tttcataggg	ttacacagaa	5520
ctaccagtt	gtgcatgtct	gatgtaattt	cacatatgaa	tg		5562

<210> 7
 <211> 1755
 <212> DNA
 <213> Mus musculus

<400> 7

tttgctttga	cagagcaagt	ttgtgggaaa	gcaactgggg	tggttttttt	tttttttttt	60
cattcaagga	agtgccacag	attaacttta	ttggccctgg	atgtatttta	tggatttttg	120
ctataaatc	caaaagcatt	gttattttatt	aaaattttcg	tgtggccact	tcattggggt	180
acacagaact	atccagctgt	gcatgtctga	tgtaatttta	catattgaat	gtatgaatta	240
cttgtcttat	tcattgtgat	acagcctcag	tccatggcgc	atccgtgtgg	ggggacccca	300
acataccag	aatcacagat	atttttccca	actattcatg	aacgtccagt	ttctttttca	360
ccacctccc	cctgtccacc	gaaagtagcc	atttcccaac	gacgtaagag	cacctccttc	420
ctggaagccc	aaactcgcca	cttccaaccc	ctgctgagga	ctgttggcca	aaatcatctt	480
ccacctggta	gcagcccaac	taactggaca	ccagaggcca	tagttatggt	gggtgctaca	540
gccaatagag	taaatagaga	gctatgtgag	atgcaggctc	aacctgtgtt	tgcaggaacc	600
cagatttaca	gtgactatag	acctggacta	gtactggcag	aagaagctca	ctattttatt	660
cctcaggaaa	cagtatatct	agctgggggtg	cattaccagg	cccagggtggc	aggacagtat	720
gagggcattt	catacaactc	accagtagct	tcaagtccta	tgaacagat	atctgaacag	780
aagccagtgc	cggggggccc	tgccctcaagt	tctgtctttg	aatttccttc	tggacaggct	840
ttcctggtag	gacaccttca	gaatttaaga	ttagattctg	gaccaagtcc	agcatcacc	900
ctctctagta	tttctgcgcc	taacagtaca	gatgctacac	atctgaaatt	tcacctgtc	960
tttgtacctc	attctgcacc	agctgtgtta	actaatagca	atgagaacag	aagcaactgt	1020
gtattttgaat	ttcatgtctc	aacaccaagt	tcttcaggag	aaggaggtgg	gattttacct	1080
caacgtgttt	accgaaatcg	gcagggttga	gtggactcaa	atcaggaaga	actgtctcct	1140
cagtcagttg	gattacattg	ccacctgcag	cctgttactg	aagaacagcg	taataaccat	1200
gccccgaat	tgaccatttc	tgtggtagaa	cctatgggac	aaatctggcc	aataggaagc	1260
ccagaatatt	ccagtgatcc	ctctcaaatt	acttcttcag	atctcagtga	ttttcaatca	1320
cctcccccta	cagggggggac	agctgcacct	tttggtctctg	acgtctcatt	gccctttatt	1380
cgctgcctc	agacagtgtt	acaagaatcc	ccactattct	tctgtttccc	ccaaggaacc	1440
acatctcagc	aggtcttatc	tgcttcatat	tcttcaggag	gatctacact	ccatccacag	1500
gttataggaa	aactttcaca	attcttttaa	actaccctat	tatacaccag	tgcattttaa	1560
ttctctatct	catttacctg	aatcatggat	ttataaggaa	tggtgtttat	tttccatagta	1620
ggattttctt	tgtggttata	aatacttgta	gtaaaaaatg	aactataaag	cagaaaatac	1680
ctggtttgcc	tgtgatttta	attgtcagaa	tcaatgacta	atctctttgg	aattttgtgg	1740
tagtaatgaa	aatgc					1755

<210> 8

<211> 433

<212> PRT

<213> Mus musculus

<400> 8

Met Tyr Glu Leu Leu Val Leu Phe Met Leu Ile Gln Pro Gln Ser Met
 1 5 10 15

Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile Phe
 20 25 30

Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Thr
 35 40 45

Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser Phe
 50 55 60

Leu Glu Ala Gln Thr Arg His Phe Gln Pro Leu Leu Arg Thr Val Gly
 65 70 75 80

Gln Asn His Leu Pro Pro Gly Ser Ser Pro Thr Asn Trp Thr Pro Glu
 85 90 95

Ala Ile Val Met Leu Gly Ala Thr Ala Asn Arg Val Asn Arg Glu Leu
 100 105 110

Cys Glu Met Gln Val Gln Pro Val Phe Glu Pro Thr Gln Ile Tyr Ser
 115 120 125

Asp Tyr Arg Pro Gly Leu Val Leu Ala Glu Glu Ala His Tyr Phe Ile
 130 135 140

Pro Gln Glu Thr Val Tyr Leu Ala Gly Val His Tyr Gln Ala Gln Val
 145 150 155 160

Ala Gly Gln Tyr Glu Gly Ile Ser Tyr Asn Ser Pro Val Leu Ser Ser
 165 170 175

Pro Met Lys Gln Ile Ser Glu Gln Lys Pro Val Pro Gly Gly Pro Ala
 180 185 190

Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val Gly
 195 200 205

His Leu Gln Asn Leu Arg Leu Asp Ser Gly Pro Ser Pro Ala Ser Pro
 210 215 220

Leu Ser Ser Ile Ser Ala Pro Asn Ser Thr Asp Ala Thr His Leu Lys
 225 230 235 240

Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr Asn
 245 250 255

Ser Asn Glu Asn Arg Ser Asn Cys Val Phe Glu Phe His Ala Gln Thr
 260 265 270

Pro Ser Ser Ser Gly Glu Gly Gly Gly Ile Leu Pro Gln Arg Val Tyr
 275 280 285

Arg Asn Arg Gln Val Ala Val Asp Ser Asn Gln Glu Glu Leu Ser Pro
 290 295 300

Gln Ser Val Gly Leu His Cys His Leu Gln Pro Val Thr Glu Glu Gln
 305 310 315 320

Arg Asn Asn His Ala Pro Glu Leu Thr Ile Ser Val Val Glu Pro Met
 325 330 335

Gly Gln Ile Trp Pro Ile Gly Ser Pro Glu Tyr Ser Ser Asp Ser Ser
 340 345 350

Gln Ile Thr Ser Ser Asp Leu Ser Asp Phe Gln Ser Pro Pro Pro Thr
 355 360 365

Gly Gly Thr Ala Ala Pro Phe Gly Ser Asp Val Ser Leu Pro Phe Ile
 370 375 380

Arg Leu Pro Gln Thr Val Leu Gln Glu Ser Pro Leu Phe Phe Cys Phe
 385 390 395 400

Pro Gln Gly Thr Thr Ser Gln Gln Val Leu Ser Ala Ser Tyr Ser Ser
 405 410 415

Gly Gly Ser Thr Leu His Pro Gln Val Ile Gly Lys Leu Ser Gln Phe
 420 425 430

Phe

<210> 9
 <211> 1716
 <212> DNA
 <213> Rattus rattus

<400> 9

gcagcttttt	ctttcattca	aggaagtgcc	acagattgac	tttattggcc	ctggatgtat	60
ttaatggatt	tttgctataa	attccaaaag	cattgttagt	tattaaaatt	ttgggtgtagc	120
cacttcatgg	ggttacacag	aactatccag	ctgtgcatgt	ctgatgtaat	tttacaatatt	180
gaatgtatga	attacttgtc	ttattcatgt	tgatacagcc	tcagtccgtg	gcgcatccgt	240
gtggggggac	cccaacatac	ccagaatcac	agatattttt	cccaactatt	catgaacgtc	300
cagtttcttt	ttcaccacct	cctacctgtc	caccgaaagt	agccatttcc	cagcgacgta	360
agagtacctc	cttctctgga	gotcaaaactc	gccacttcca	acccctgctg	aggactgttg	420
gccaaaatca	tcttccacct	ggtggcagcc	caactaactg	gacaccagag	gccatagtta	480
tggtgggtac	tacagccaat	agagtaaata	gagagctatg	tgagatgcag	gtccaacctg	540
tgtttgagac	aaccagatt	tacagtgaat	atagacctgg	actagtactg	gcagaagaag	600
ctcgctattt	tatcctcagg	gaaacagtat	atctaggttg	ggtgcattac	catccccatg	660
cggcaggaca	gtatgagggt	atttcataca	actcaccagt	actgtcaagt	cctatgaaac	720
agataactga	acagaagcca	gtgcctgggt	gccctgcctc	aagttctgtc	tttgaatttc	780
cttctggaca	ggctttccta	gtaggacacc	ttcagaattt	aagattagat	tctggaccaa	840
gtccagcatc	acccctgtct	agtattttctg	cgcctaacag	tacagatgct	acacatttga	900
aatttcaccc	tgtctttgta	cctcattctg	caccagctgt	gttaactcat	agcaatgaga	960
acagaagcaa	ctgtgtatgt	gaatttcattg	ctcaaacacc	aagttcctct	tcaggagaag	1020
gaggtgggat	tttacctcaa	cgtgtttacc	gaaatcgaca	ggttgacgtg	gactcaagtc	1080
aggaagaact	gtctcctcag	tcagttggat	tacactgcca	cctgcagcct	gttactgaag	1140
aacagcgtaa	taaccatacc	ccagaattga	ccatttctgt	ggtagaacct	atgggacaaa	1200
actggccagt	aggaagccca	gaatattcca	gtgattcctc	tcaaattact	tcttcagata	1260
tcagtgaatt	tcaatcacct	ccccctacag	ggggaacagc	tgacaccttt	ggctctgacg	1320
tctcattggc	ctatattcgc	ctgcctcaga	cagtgttaca	agaatcgcca	ctattcttct	1380
gtttcccca	aggaaccaca	tctcagcagg	tcttatctgc	ttcatattct	tcaggaggat	1440
ctgcactcca	tcacaggtt	ataggaaaac	tttcacaatt	cttttaaaat	accctattat	1500
gcaccaatgc	atttaaattc	tctgtctcat	ttccctgaat	catggattta	taagaaatgt	1560

tattttctta gtaggattttttgtgggtt ataaatatat gtactaaaaa aaagaacttt 1620
 aaagtggaaa atacctggtt tacctgtgat tttagtgtgc agaataatg actaatatct 1680
 gtggaatttt gtggtagtag taaaaatgca attgtc 1716

<210> 10
 <211> 434
 <212> PRT
 <213> Rattus rattus

<400> 10

Met Tyr Glu Leu Leu Val Leu Phe Met Leu Ile Gln Pro Gln Ser Val
 1 5 10 15
 Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile Phe
 20 25 30
 Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Pro Thr
 35 40 45
 Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser Phe
 50 55 60
 Leu Glu Ala Gln Thr Arg His Phe Gln Pro Leu Leu Arg Thr Val Gly
 65 70 75 80
 Gln Asn His Leu Pro Pro Gly Gly Ser Pro Thr Asn Trp Thr Pro Glu
 85 90 95
 Ala Ile Val Met Leu Gly Thr Thr Ala Asn Arg Val Asn Arg Glu Leu
 100 105 110
 Cys Glu Met Gln Val Gln Pro Val Phe Glu Thr Thr Gln Ile Tyr Ser
 115 120 125
 Asp Tyr Arg Pro Gly Leu Val Leu Ala Glu Glu Ala Arg Tyr Phe Ile
 130 135 140
 Leu Arg Glu Thr Val Tyr Leu Gly Gly Val His Tyr His Pro His Ala
 145 150 155 160
 Ala Gly Gln Tyr Glu Gly Ile Ser Tyr Asn Ser Pro Val Leu Ser Ser
 165 170 175
 Pro Met Lys Gln Ile Thr Glu Gln Lys Pro Val Pro Gly Cys Pro Ala
 180 185 190
 Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val Gly
 195 200 205
 His Leu Gln Asn Leu Arg Leu Asp Ser Gly Pro Ser Pro Ala Ser Pro
 210 215 220
 Leu Ser Ser Ile Ser Ala Pro Asn Ser Thr Asp Ala Thr His Leu Lys
 225 230 235 240
 Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr His
 245 250 255
 Ser Asn Glu Asn Arg Ser Asn Cys Val Phe Glu Phe His Ala Gln Thr
 260 265 270

Pro Ser Ser Ser Ser Gly Glu Gly Gly Gly Ile Leu Pro Gln Arg Val
 275 280 285

Tyr Arg Asn Arg Gln Val Ala Val Asp Ser Ser Gln Glu Glu Leu Ser
 290 295 300

Pro Gln Ser Val Gly Leu His Cys His Leu Gln Pro Val Thr Glu Glu
 305 310 315 320

Gln Arg Asn Asn His Thr Pro Glu Leu Thr Ile Ser Val Val Glu Pro
 325 330 335

Met Gly Gln Asn Trp Pro Val Gly Ser Pro Glu Tyr Ser Ser Asp Ser
 340 345 350

Ser Gln Ile Thr Ser Ser Asp Ile Ser Asp Phe Gln Ser Pro Pro Pro
 355 360 365

Thr Gly Gly Thr Ala Ala Pro Phe Gly Ser Asp Val Ser Leu Pro Tyr
 370 375 380

Ile Arg Leu Pro Gln Thr Val Leu Gln Glu Ser Pro Leu Phe Phe Cys
 385 390 395 400

Phe Pro Gln Gly Thr Thr Ser Gln Gln Val Leu Ser Ala Ser Tyr Ser
 405 410 415

Ser Gly Gly Ser Ala Leu His Pro Gln Val Ile Gly Lys Leu Ser Gln
 420 425 430

Phe Phe

<210> 11
 <211> 486
 <212> PRT
 <213> Takifugu rubripes

<400> 11
 Ser His Phe Glu Cys His Ser Ile His Cys Ile Ser Leu Ser Gln Gln
 1 5 10 15

His Ser Val Pro His Pro Tyr Ser Gly Gly Ala Gly Gly Gly Ala Gly
 20 25 30

Gly Asn Ile Pro Leu Ala Ser Leu Pro Asp Pro Thr Thr Val Leu Phe
 35 40 45

Pro Ser Ile Pro Glu Arg Pro Ile Ser Phe Ser Pro Pro Pro Thr Met
 50 55 60

Pro Pro Lys Ala Tyr Asn Asn Gln Arg Arg Lys Ser Thr Ser Ile Leu
 65 70 75 80

Glu Ala His Thr Arg His Phe Gln Pro Ala Tyr Thr Arg Tyr Gly Ser
 85 90 95

Asn Leu His Pro Phe Ser Gly Met Glu Thr Val Glu Pro Pro Pro Leu
 100 105 110

Phe Met Val Asn Pro Gly Phe Ala Ala Ala Ala Gln Arg Leu Gly Val
 115 120 125
 Gly Glu Pro Leu His Leu Pro Gly Gln Ser Asp Pro Ser Leu Tyr Gly
 130 135 140
 Tyr Lys Asp Met Arg Ala Glu His Gly Glu Ala Met Arg Arg Leu Ser
 145 150 155 160
 Leu Asn Gln Ala Ala Leu Leu Asp His Tyr Glu Ala Met Ala Tyr Gly
 165 170 175
 Gly Tyr Pro Met Thr Ala His Gln Leu Ser Arg Leu Ser Phe His Gln
 180 185 190
 Gln Arg His Ala Ala Ala Ala Ala Ala Ser Phe Ala Phe Glu Gln Pro
 195 200 205
 Gly Phe Leu His Pro His Leu Ile Gln Arg Met Ser Ala His Ser Pro
 210 215 220
 Val Pro Pro Pro Leu Ala Pro Leu Ser Ala Pro Ser Gly Thr Ser Met
 225 230 235 240
 Ser Ala Ser Ser Glu Gly Cys Tyr Pro Ser Pro Gln His Ala Ala Val
 245 250 255
 Ser Thr Phe Ala Ile Ser Ala Pro Pro Val Met Pro Asp Ala Ala Pro
 260 265 270
 Pro Thr Gly Asn Val Phe Glu Phe His Leu Ala Ala Ala Gly Asp Pro
 275 280 285
 Asn Leu Met Ala Ser Arg Leu Phe Arg Ala Arg Arg Ser Ser Met Asp
 290 295 300
 Leu Pro Leu Glu Asp Asn Thr Gly Ala Gly Ser Gly Gly Ser Gly Thr
 305 310 315 320
 Tyr Val Arg Leu Gln Pro Val Thr Glu Glu Leu Tyr Ser Tyr Val Ser
 325 330 335
 Pro Glu Leu Pro Leu Pro Pro Gly Ser Leu Phe Leu His His Ala Gly
 340 345 350
 Leu Thr Ala Lys Asp Lys Ser Pro Asp Pro Ser Ser Asp Ser Leu Ala
 355 360 365
 Ser Ser Asp Asn Gly Glu Phe Gln Ser Pro Pro Pro Pro Pro Leu Leu
 370 375 380
 Pro Ser Phe Asp Ser Ser Ser Ala Gln Ser Ile Pro His Thr Ala Ser
 385 390 395 400
 Thr Ala Met Tyr Asp Pro Ser Thr Gly Met Leu Pro Ile Asp Pro Gln
 405 410 415
 Gly His Pro Pro Ser Met Gln Ser Phe Leu Pro Pro Thr Pro Ser Ser
 420 425 430

Glu Leu Pro Leu Gly Ala Pro Ser Thr Gln Leu Glu Thr Met
 435 440 445

Gln Ser Ala Trp Ala Arg His Gly Gly Val Val Pro Ala Gln Pro Asp
 450 455 460

Met Thr Tyr His Glu Ser Leu Leu Ala Met Gln Ala Val Asn Pro Ser
 465 470 475 480

Leu Val Leu Leu Arg Phe
 485

<210> 12
 <211> 13546
 <212> DNA
 <213> Homo sapiens

<400> 12

ccctcaagta	tctacaaaga	tggttagggaa	atgcataatg	cagatttttac	agttcgggtta	60
aggccagtc	tgagagaatt	ctttcttcct	cctcttctct	acttttttta	caatgccttt	120
tttttttttg	gcgggggggtg	gtgggtggggg	gtgttggttt	ctttcaatat	actactgctt	180
aattttaccct	tttattctgt	aggttagagtc	tgggtatgtc	tgtgaagggtg	atcacaaagac	240
catgggctaaa	gctatcaaag	acagagtatc	attaattaag	aggaaacgag	agcagcggca	300
gttggtacgg	gaggagcaag	aaaaaaaaaa	gcaggaagag	agcagtctca	aacagcaggt	360
agaacaatcc	agtgtctccc	agacaggaat	caagcagctc	ccttctgcta	gcaccggcat	420
acctactgct	tctaccactt	cagcttcagt	ttctacacaa	gtagaacctg	aagaacctga	480
ggcagatcaa	catcaacaac	tacagtacca	gcaacccagt	atatctgtgt	tatgtacgta	540
tcttggaag	tggacagata	ggctaattgt	tctcctaatt	ggttttttta	gcttcttggt	600
atthtgcact	ggcttttttg	ttttgaatta	tgaatcctac	tattatacaa	aatttgaaat	660
gggtgtcgtc	tagcttctct	ttattcttat	atctgtgact	aggagaatag	aaggatgggg	720
tggttaaggaa	atgtgattag	gatacaaaat	aattttataa	gcttcaacat	cttggagggtt	780
tcaatttgta	gatttctact	tttcacaatt	tatttaataa	aaacagattt	aataatgaag	840
aatgaacaag	cttggctact	aaaattgaac	ccatcagctc	ggttgagctc	ttttagcaaa	900
gtgaatgggt	atatagacta	aaaactaatg	agaacactgg	ttattaattt	aattacttgt	960
ttaataactga	gaacctcctg	gcctccataa	ttaagacttt	gaaagctgtg	tcaaatttaa	1020
ggcttatagt	gtttcatgct	gtttccctta	ctcctcaggt	ggtaagattc	tgtgatagcc	1080
atatggattc	ctcattttaa	tatagaaatc	tactgaatat	ggctacaatt	cattttttat	1140
ttagttcaaa	tctgtgcaag	gaataactag	ttacccttaa	tataatgggt	ctaaatatga	1200
gaaaatgtga	acctccattt	tgtgatttgt	ctttctctct	ctcttttttt	tggcgattca	1260
tttttccttc	agctgatggg	acggttgaca	gtggtcaggg	atcctctgtc	ttcacagaat	1320
ctcgagttag	cagccaacag	acagtttcat	atggttccca	acatgaacag	gcacattcta	1380
caggcacagt	cccaggcat	ataccttcta	ctgtccaagc	acagtctcag	ccccatgggg	1440
tatatccacc	ctcaagtgtg	gtaagtaaat	gcttaagagc	atgtaatact	acaatgagag	1500
ccaatggata	gtctgctaaa	ttaaaattct	ttgtacaaac	caagtaacag	tatgagctg	1560
aagtagaaaa	tataaaatgc	tgtgtctcta	aagccttatt	atagagtagg	atatagacat	1620
taatgtgtgg	gcattagctg	aactgcctac	atgaagcagc	tatgtgcaag	atthtagatta	1680
gtctgcattt	ttactaaagg	gaacgctctc	tttagcactt	aagatgttga	aactattcac	1740
agcaacaaaa	gctggacata	caaatgtttt	gcgggtgggg	cgagaagagt	ttttcaaaag	1800
cttggttaaaa	tgattactaa	gtatttcctg	taagattccc	tattgtacaa	aacacaggta	1860
agaccttcaa	accattatag	aatgctgcat	gtgagtactc	tagattgagg	gaatccttta	1920
tttgaatgac	tgtttaattt	agcgcttggt	ttgagttaat	catcagttgc	tgtgtgtggt	1980
cctgcctttt	aaagtctgtg	ttggacacaa	gttccttccc	cagtatatgt	tcttttgctt	2040
tttcccttta	gttggttaact	gttttattac	tgatttattt	tcagaatcaa	cttatagaaa	2100
ttatttcctg	aaacatttta	tgcttggtata	cacatacaca	cacacaatta	tcctgtgttt	2160
tacattagac	actcttggtt	ctcatataga	gaatgttaaa	gtagatcatc	agataatacc	2220
ccattttctt	tggtgacctc	agaggaaaat	aaaatacgtc	atctgcctct	gaaagggtatt	2280
agaaaataag	taactgtgca	ctgtacagca	cttggttagt	gagatcttta	agcagacaga	2340
ttttccctt	ctcagaaagc	ttttaatcta	aactttgtgg	aattcttgct	cttcttttaa	2400
tgaaaatctt	tctotgttaa	gaaaataactt	ctgcttgctt	gattttttat	cctacagctt	2460

ggaataaaag	atgttcac	gtttcagagt	ccgtttaatc	ctgctttggc	taggtctcaa	2520
aatgaatggc	cacagtaaca	caaataaaaa	actaatatac	taacttttta	aaacaatcta	2580
ctttctgtct	ttaaataaat	tgagggtaga	atagaagaac	tttataaaga	ctgcacagaa	2640
aatatatttc	tgtgttctga	caaataattac	ttgtcagcat	ttcttgtcag	tttctacctt	2700
taattgtttt	ttttctggat	tattataact	agattttgct	aagacttttt	ttttaaaatt	2760
ttctcttgtc	tatactcatc	atggggggagc	attaggtctc	tatttttagaa	agcaagtaga	2820
atgacatgaa	gatattatit	gagccattaa	tctagctata	tttattgatg	ctctaaaagc	2880
agacgtttct	gctttttaat	gacagtattc	tatggaggta	ttatgaagaa	ttattaaata	2940
ttgagaatit	aaaaatttgt	ctaattctat	tcttgtctca	tttcagggaat	gtggcctaca	3000
ataatatgga	caaattataa	gtgaaatit	agaaattaga	ataaaaagaaa	taaggcttaga	3060
gagaaatcaa	aacaggaatg	aaactaaaca	atcatatgat	gatatgtaaa	tacttgggtat	3120
tggtggggcca	cagatttggc	ccaaagcatt	ctacacagcc	aagttttgga	aatgaaaaca	3180
tagtttcata	gttcacaaca	tctaaaagga	aaataacagt	cttattcaag	acaaccttta	3240
tggtattgat	catactgcct	tatagccata	ttaaaccctg	tgccctgaaca	tttttttaag	3300
tttttttttt	tttttttgac	agcgtctcac	tctgtctccc	aggtgggaat	acagtggcgt	3360
gatcttggct	cactgcaact	ttcacctcct	gggttcaagt	gattctcctg	cctcagcctc	3420
cccagtagct	aagattacag	gcacccatca	ccacgcccag	ctaatttttg	tatttttttt	3480
agtagagaca	cagtttcacc	gtgttggcta	ggctggtctt	gaactcttga	cctcaagtga	3540
tccacctgcc	tccgctccc	aaagtgccag	gattacaggt	gtgagccaac	atgcccggcc	3600
taaattatta	tttagagatt	cagattgaag	atattctccc	attgattact	ttgtaaaagt	3660
acaaaatatt	tcctatatcc	agttgtgaat	ttcttctaga	aagctaaact	ttttcctagt	3720
agatcatgac	ttttttacct	ttacaaataa	agagtggcaa	gtataagaaa	tactgtgagc	3780
caacaacttt	ccatctcatg	tcttctaaaa	tagaaatgag	aatctttttc	tgaggtttcc	3840
agagctcttg	caaaaataatc	caatctttaa	attctctgta	tcttttaagt	tactttaaac	3900
taatagaatt	cattcttgtt	gaatcttaag	tttgccaaag	ttgcaagcta	atattcattg	3960
ttttccctgt	ctgcaccctt	cagtcagttg	gtcattctta	aattcataca	cttgtaaacac	4020
ttggtgttat	ttcagaatat	ttagcatgat	aactgatcat	tatcatgatc	attcattaaa	4080
cctgtccttg	actgacaacg	cagaatagtt	gtgaatagtt	gtgatctgtc	agatgtgcta	4140
tgttgagtta	tagctctcct	caaaaaaaaa	actgactttg	tggaattggg	agaggatgga	4200
acatttcttc	atgcaactat	gccctgttat	cattgtactt	ttgtttattt	tgttttcaca	4260
gtactgtgtt	tttcatgtgt	gtgtttgttt	tgtgttgagc	ctcgtcgtgg	ccgtagcatg	4320
tcggtttgtg	ttcccatctt	tctgtctgtg	tctttctcct	ccgcatctct	cccagtgtct	4380
ttccacccca	ccgccagtac	tgtctgcacc	tccctcctcc	tccctcctcc	ggactgcccc	4440
gaggaaactt	ttgccgaaaa	gcttttctaaa	gcattggaga	gtgtcctgcc	tatgcaactc	4500
gcctctcagc	gcaagcaccg	acgctccagc	ctgccttccc	tctttgtcag	tactgtatgt	4560
aactgtaaac	ttctgacaaa	tgaacaatta	ttaccaatga	atacatccag	gcaacaagaa	4620
ttttaattaa	aattgaataa	agaacaggac	taaaacttgt	tatattatgc	ttctaggata	4680
ccagttattc	ctgtaggaaa	cttttttgtg	ataaggcttt	aagggggaat	aaaaagcaag	4740
gaggtaatga	gagagaatac	ctataatgtc	cttggcacat	accagcatgt	gtccaataaa	4800
atttaccttt	cattatatct	tcctgtttcc	tatcgacata	aaattctttc	attttcatat	4860
tctagtttta	aaaggattaa	ttttaaaagt	attgccttaa	tcaaaagcaa	attttatcaa	4920
acattaccca	aacatttttt	ctcctcactt	tattctaagg	agacactaat	aaaactggat	4980
gagtcctttt	tttttttaat	aaccttagtc	agcaaaagaat	aaccttggag	tgaacattta	5040
ttttaattta	atcattttatg	taagatgatg	tatcagcaaa	ccaaagaact	acataatgtt	5100
cccccatgga	agtatttttt	ccttaggttt	caacctaaag	tgggactgta	gctctaattga	5160
atgcattaga	tgcataatata	gtatgtatga	ctttcttttg	ctccattaat	gtttataaaa	5220
gatctatctt	ataaaaattat	tttaagcatg	tggttccttg	ctgcttttcc	aaataaagaa	5280
tatagttaac	tttctgaact	agcaacccca	gaattttcag	atttttactt	aggagaatga	5340
agaactatac	agtgttatta	tggcttttgag	acataccatg	agtaaataac	aataaaatga	5400
gcattctgac	tgaaaaataga	aatccttatt	tagctattgt	gaacagttga	tctactttgc	5460
cagctcgata	aggatcttcc	ctatggaaga	atgaaactga	attcttttaa	aaagcggggg	5520
aggctgtgat	agtgaattct	tactattatt	tacttattta	tttatgtttt	ggagatggaa	5580
tctcgctctg	tcgcccaggct	ggagtgcagt	ggcacgatct	cgactcacgg	caacctctgc	5640
ctcctgggtt	caagtgtttc	tcctgcctca	gcctcccag	tagctaggac	tacaggcgca	5700
tgccaccaca	cccagcta	ttttgtattt	tagtagagat	gggtttcacc	atgtttggcc	5760
aggatgggtct	tgatctcctg	acattgtgat	ccacccctt	cagcctccca	cagtgtctgg	5820
attacaggcg	tgagccaccg	cgcccagegg	gtaatgaatt	attatagtgt	atattgtttg	5880
catttggcaa	gctgtgtttc	taaatttagt	cttaaatcta	gaactgcttc	tagaataaaa	5940
agtgtcatca	aataaaaattg	gctgtcattt	atattaccct	tttcagttga	ttagaacacc	6000
aggatgtcag	aatcagaaaa	ttcatggtta	cggccgggct	cagtggctca	cgctgttaat	6060
cctagcactt	tgggaggccg	aagttagcag	atcacctgag	gtcaggagtt	cgagaccagc	6120

ctggccaaca	tggcgaa	tgtctctact	aaaaatacaa	aaattagctg	gggtgggtgg	6180
cgcacgcgtg	tagtcccagc	tacttggggag	gctgaggcag	aagaatccct	tgaaccagg	6240
aggcggaggt	tgcagtaagc	caagattgtc	actgcactcc	agcctgggcg	aaagaggaag	6300
actccatctc	aaaaaaaaa	aaaaagaaat	ttcatgggta	tgcactctt	atztatgatc	6360
agaaaaatgg	acattttgtg	atttaactct	gtaacatggt	tcatgtagta	aaaatataat	6420
aaaactatta	atcatctagc	ttgggagaga	taggaagaa	acattactgt	cactagtcaa	6480
attatatatc	ttttactatc	cacaaaaaat	ctcttctgat	ttctgggttag	aaggcatact	6540
attaattgat	aagaaaataa	aactgaaggc	ctctaacata	tcacagggtg	ataagaatat	6600
agggaaggtt	agttcaatag	tttaaattaa	agcacacttc	ttacagtata	gaactagtcg	6660
ggcttttatg	ccttggttta	gttcttactc	ttcctttaac	tctttttctg	ttgatgtaat	6720
ttacattaat	gcttaagagt	gaacttttta	agtgtgggta	aaaacgaaat	aattacttac	6780
aaagtttaat	tcttccattt	cctttgagag	aggaaaggtt	tggaaaagca	gctcttatct	6840
aaagcaaaga	gccacagat	tgatttcatt	ggccctggat	gtatttaatt	gatttttact	6900
atgcacataa	tttccagaag	cattgttatt	tatttattaa	ttataaattt	agtgtaacca	6960
tttcataggg	ttacacagaa	ctaccaggtt	gtgcattgtc	gatgtaattt	cacatatgaa	7020
tgtatgaatt	acttgtctta	ttcatgttga	taagcctca	gtccatggcg	catccgtgtg	7080
gggggacccc	aacataccca	gaatcacaga	tatttttccc	aactattcat	gaacgtccag	7140
tttctttttc	accacctccc	acctgccac	cgaagtagc	catttcccag	cggcgtaaga	7200
gcacctcctt	cctggaagcc	caaactcacc	acttccaacc	cctgctgagg	actgttggcc	7260
aaagtcttct	tccacctggt	ggcagcccaa	ctaactggac	accagaggcc	gtagttatgt	7320
tgggtactac	agccagtaga	gtaactggag	agtcatgtga	gatacagggt	catcctatgt	7380
ttgaaccatc	tcaagtttac	agtgaactata	gacctggact	agtacttcca	gaagaagctc	7440
actattttat	tcctcaggaa	gcagtgtatg	tagctggggt	acattaccag	gcccggtgg	7500
cagaacagta	tgagggcatt	ccatacaact	catcagtact	gtcaagtcct	atgaaacaga	7560
tacctgaaca	gaagccagta	caagggggcc	ctacttcaag	ttctgtcttt	gaatttccat	7620
ctggacaggc	tttctgtgta	ggacaccttc	agaatctaag	attagattct	ggattgggtc	7680
cgggatctcc	cctctctagt	atttctgcac	ctatcagtac	agatgctaca	cgtttgaaat	7740
ttcacccctg	ctttgttctt	cattctgcgc	ctgctgtgtt	aactcataac	aatgagagca	7800
gaagcaactg	tgtatttgaa	tttcatgttc	acacaccaag	ctcctcttca	ggagaaggag	7860
gtggaatttt	acctcagcgt	gtttaccgaa	atcggcaggt	tgcagtggac	ttgaatcaag	7920
aagaactgcc	tcctcaatca	gttggattac	atggctactt	gcagcctgtg	actgaagaaa	7980
agcataatta	ccatgcccc	gaattgaccg	tttctgtggt	agagcctatc	ggacagaact	8040
ggccaatagg	aagcccagaa	tattccagtg	attcctcaca	aatcacttct	tcagacccca	8100
gtgattttca	gtcacctccc	cctacagggg	gagcagctgc	accttttggc	tctgacgtct	8160
caatgcctct	tatccatctg	cctcagacag	tgttacaaga	atccccactt	ttcttctgtt	8220
tcccccaagg	aaccacatct	cagcaggtct	taactgcctc	attttcttca	ggaggatctg	8280
cacttcatcc	acaggttata	ggaaaaacttc	cacaattatt	ttaaactacc	ctactttgca	8340
ccataacatt	taaattttct	attccttatt	tccctgaatc	atggattttg	gagaaatatt	8400
gtttaatttt	atcagtagag	tttccccatc	tttggggggt	tgtgaactac	atatatgcat	8460
ttaaaaacaa	aatgtgagag	aagctacctg	atttacctat	tatatgtgaa	aaccagtgga	8520
aaaaacacaa	aaactagaat	tttagtcatt	cttcacaatc	acgacttcta	tgcacgttat	8580
tttcaaccag	tagtgaaaat	gcaagtgtat	gtaatgtatg	gttgacccag	cattatttag	8640
gaatacaaat	cttaagtatt	actttcttcc	tccaaacaag	tttttaaaaa	ataggataaa	8700
ttttttttct	ataaaatata	aaacatggaa	aatagggaa	gctgtttttg	aggtaatatt	8760
aataatacac	agaattttca	ttagtgtcga	aggatctaaa	aagacaaaagt	atatcatggg	8820
aataaaaaaa	gatagaaaag	gaaacagttt	aggaatttgc	cttaacaaat	gaaaatgcct	8880
ttttaaaatg	gcatcagtc	agcaagttgc	tgtgcattat	tatatgtcca	aataaaatgc	8940
taattcataa	aattaaggac	tgattctaat	ttttagcaac	taacaatcta	gtgagtgggt	9000
ggcttttttt	ttttgttata	agaattttgt	atatgttgca	agctaaggaa	cagtagacat	9060
cttaggcctt	gtattttcct	caacaaattt	aatttttagtt	atztatgtgt	atgagcatat	9120
acacacttag	caaataaaga	caatgtaatg	aatgtaatag	aattcaaaag	tacagtcagt	9180
gtttactgtt	tcatatattc	atttgggtata	ttttatcaaa	tgtgtacatc	ttataaaatc	9240
tacatagaag	actatgcagt	tatatattga	agacatttgg	atatcctaca	taaagtacta	9300
tgtgttggtc	ttcctacttt	cttcctatgt	tcttccctca	tcaggagcat	aaacaactcc	9360
caaaatgtgc	acaaattatc	aattatttag	tgctgtacta	agacacaaaa	agatccagaa	9420
gcttgacagc	tatacatggg	catgcagtat	attcacatca	gtattcaagt	tgcagcccct	9480
attataaata	ttaatgtatt	gccagtaatt	atcctttaga	ctttccaaag	taaaaaatac	9540
caggtagtta	atttatagac	atactacaga	aggaatcaat	atztatgata	gttaagctgc	9600
agtggaatgt	gatattatat	ccatgtaagc	tttaaggaac	cccttgtaaa	ctgaaatgga	9660
atgacaaaag	tggtagcctt	ttcatttctg	taacatacct	atatatagtc	ctgcatcatt	9720
tctcactttc	tagtgagaaa	cactgtagtt	ggggtttctc	agactcaaca	ctgttgacat	9780

ctgaggttgg	ataattct	gttgtggtgg	gctgttctat	gtattgtagg	ctattagaa	9840
catccttggg	ttatacctcc	taaatactag	tagcagcaat	ctccctcagt	tgtgacaacc	9900
agaaatgtct	ccacatattg	ccaagtgcc	ccctgggggt	ggcaggattg	ctcctggata	9960
gcaatgatcg	ccctaggtac	ttggcagttg	atgataacag	atTTTTtgat	acttttgata	10020
aaagtcattg	tggcactaca	agaagtatta	ccaaattagc	attaccaa	acdtttcttt	10080
aattcttcag	ataaatttca	gattgatagt	atttaacttt	tggtcctgat	tattgtattt	10140
ctaataacaa	atTTTTtatt	gaatctgacg	attctagcca	gctaaaatta	cagcctgagc	10200
agcattgaca	gcttaggttt	tactgtcag	tgggttttgc	atgggttggg	atcagaacat	10260
gttgtgtggg	atgatccttt	tctaaatttt	tcttaca	gttgaccag	aggcctctcc	10320
catacataat	cagggttaatg	tttcattttg	atcaagcaaa	ctctgaggag	attaaatatt	10380
cattgcaaaa	gcctgacctc	tatacactta	ctttaggcct	tctctaattt	gttgtgttca	10440
cttctttcct	tcaggcacag	gggcagagcc	agggtcagcc	atcctcaagt	agcttaacag	10500
gggtttcatc	ttcccaacc	atacaacatc	ctcagcaggt	gagaacaatg	cattgcaaca	10560
ttttatgaat	tagcagtggc	cagaacacta	ttttaacttt	tgtgatTTgt	tatataaatg	10620
gcctgctaag	tttttgaaag	gaaaatttaag	aggcatacca	tgtcttTgt	tcagtaacaa	10680
agaaaaaa	gagacttttc	tgtgtttcag	aaaaatctct	tttgtttaat	atacttttga	10740
acaatttagt	agacatacat	tattactgat	ataattatct	ttttaaaaga	ggattctaca	10800
cctttttttt	gagacaaagt	ctcactctgt	caccaggt	ggagtgcagt	ggtgcaatct	10860
tagctcactg	caaccaccgc	ctcccaggtt	caagcagtc	tcctgcctca	gctttccacg	10920
tagctgggac	tacaggtgcc	tgccaccaca	ctcagcta	ttttgtattt	ttagtagaga	10980
cggggtttc	ccatgtggc	caggctggc	tcaaactcct	gacctcaggt	aatccaccct	11040
cctcggcttc	ccaaagtgc	gggattacag	gcatgcgcca	cctcaccag	cctggattct	11100
acacataatt	tacggtacaa	agagtatgtc	attccttttc	tctttttttc	cactgcttgc	11160
agcagtgatt	tgtcattacc	ttcattgtgc	ttgcctctac	taatcctgca	cttcttttta	11220
attatttcta	atctgtagaa	tgatgtgtca	gattttacgc	taagaagcat	ctaataaaaa	11280
cctaagaaca	attacttcaa	gaacatttat	ttattttatt	atTTTTtatt	atTTtgagac	11340
agccttgctc	tgttgcccag	gctggagtgc	agtgggtcaa	tcttggtcca	ctgcaacttc	11400
catctcccag	gttcaageta	ttctcacgcc	tcagccgccc	gagtagctgg	gactacaggc	11460
gcgtgccact	atgccagct	aatttttTga	ttacgttcta	cttatgcctg	gctaatttag	11520
tagagatggc	gtttcaccat	gttggccagg	ctggcttTga	actcctgacc	tcaagtgate	11580
caccgcgttt	ggcctcccaa	agtgcgtggga	gccactgtgc	ccagccaaaa	tattttaact	11640
tttgttttgg	ggggcttTgt	ccagaaagtt	caaggaattc	aagagattag	tgtcatcctg	11700
atacatccaa	cagtaaaact	ggtaattcta	aaattaaata	tttatagtgc	atcactgata	11760
ttagttaata	gtaattattt	gcaagggaat	aatgcatttt	tggcataatt	gaatacttat	11820
tagctattct	acataaaaatt	tgtagccact	atgtatgtaa	tgcaaaacttc	ctaattgtag	11880
ttcccact	gggtttttta	atcagcaaac	tctaaactgg	cttctgtttc	tattagtttt	11940
tggttttgaa	acaaagtttt	cggttcttaa	ttacaaatgg	ccctttttca	gtccaacata	12000
tggatcatgta	gtcattcgat	actggtttaa	gatcaggagt	tgcttgata	agatgaaaat	12060
cacttaaagg	cactttctac	agaaacagtt	gattagtttg	tttatacatg	cattcactcc	12120
gtgaactgaa	ggactgggac	ttatatgctt	tgttttgcac	ctgtcctttc	ttccaaacaa	12180
gactggttaa	attcaagtaa	acatattttt	gccaacctta	acaaaatctg	ttttattcta	12240
ttaacagttc	ctaattagaa	tgcaaattac	aagtcacata	aagctgggtg	agaatggaag	12300
tactgagtac	taatctccta	acttgcttat	ctttggatgt	agacgaacac	caatgaactg	12360
agcaaattgt	atgtgctctg	ttttttcttt	tccaaaggaa	ttgggttatct	gtcctgtatt	12420
tgggatgatg	tatgtatact	tactatcaaa	tgaagaagat	aataaaataa	ttcattgagt	12480
ttcatgttaa	gaaagagtgt	gagaaaagaa	gagagcagta	ccataaatgt	cagattaaaa	12540
gaattaggtg	gctacctctc	caggagaaaa	aaatatggct	aaatgctatc	tctacagcag	12600
aaacaaaaat	taatttttTgt	ttcagtgatg	tttttTgttt	agatactttt	atagcatgtg	12660
ttgtagaatc	acaaaattga	gccctcgata	tatataggaa	gtgtattcgg	tctagaataa	12720
ttaactttctc	caaagtcact	tagataataa	gggagctaca	aattataacc	cagaattcaa	12780
gattcctaga	atagtattgt	ttcttccatt	ttaaataaaag	ttctgcaggg	acttaattgat	12840
tctgcaaatg	tacctcagaa	gctgtcttta	tttgttttaa	ttattggagt	tccacgtttg	12900
aaaaaagaaa	ggttgaaaac	tttttTgtact	atatggatag	tctattttta	gacatgattt	12960
ttgtttttacg	tactttcagg	tcccaatccc	accttagaga	accatatctc	ttgttttatt	13020
ttagtgagga	aagaaatata	cttttgctaa	aatTTacttt	atTTcatgctg	tttcttgaac	13080
atagtaaagt	aggaattgct	tttttTgtgtg	tttttTgtagg	gtctcgtata	gtgatatact	13140
agatcttagc	aaggtaccaa	aaatagctgc	ctcctgctga	aaagctttgt	gaccacaaac	13200
aaagttgtga	actttgtcag	acacagtgcc	agotacttgg	aaggctgagg	caggaagatc	13260
acttgagccc	aggagttcaa	ggctacagtg	aactatgata	ttgccaccat	actccagcct	13320
gggcaacaga	gcatgactca	tctcaaaaaa	aaaaaagcac	cgaagtTgtg	aactacctac	13380
cctactatcc	ataaagtctg	tttgaactgg	tttcatgagg	caagcttggg	cagagcatga	13440

tgctaaactg ttaagcta ccacccctgc ctggctgaag tgctgattat tagccagtgg 13500
 gactgcgcac agcagatcat catgtacttc tggcatatcc agaggt 13546

<210> 13
 <211> 509
 <212> PRT
 <213> Tetraodon nigroviridis

<220>
 <223> Partial amino acid sequence

<220>
 <221> misc_feature
 <222> (326)..(397)
 <223> X = undetermined amino acid

<400> 13
 Leu Cys His Ala Leu Ser Phe Ile Ser Ser Gly Leu Cys Trp Gly Phe
 1 5 10 15
 Ser Leu Phe His Ser Glu Cys His Ser Glu Ile His Cys Ile Ser Leu
 20 25 30
 Ser Gln Gln His Ser Met Pro His Pro Tyr Ser Gly Gly Ala Gly Gly
 35 40 45
 Gly Ala Gly Gly Asn Val Pro Leu Ala Ser Leu Pro Asp Pro Thr Ala
 50 55 60
 Val Leu Phe Pro Ser Ile Pro Glu Arg Pro Ile Ser Phe Ser Pro Pro
 65 70 75 80
 Pro Thr Met Pro Pro Lys Ala Tyr Asn Thr Gln Arg Arg Lys Ser Thr
 85 90 95
 Ser Ile Leu Glu Ala His Thr Arg His Phe Gln Pro Ala Tyr Pro Arg
 100 105 110
 Tyr Gly Ser Asn Leu His Pro Phe Ser Gly Met Glu Ala Val Glu Pro
 115 120 125
 Pro Pro Leu Phe Met Val Asn Pro Gly Phe Pro Ser Ala Ala Gln Arg
 130 135 140
 Leu Gly Val Gly Glu Pro Leu His Leu Pro Gly Gln Ser Asp Pro Ser
 145 150 155 160
 Leu Tyr Gly Tyr Lys Asp Met Arg Ala Glu His Gly Glu Ala Val Arg
 165 170 175
 Arg Leu Ser Gln Ala Ala Leu Leu Asp His Tyr Glu Ala Met Ala Tyr
 180 185 190
 Gly Gly Tyr Pro Met Ala Ala His Gln Leu Gly Arg Leu Thr Phe His
 195 200 205
 Gln Gln Arg His Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala
 210 215 220
 Ser Phe Gly Phe Asp Gln Pro Gly Phe Leu His Pro His Leu Val Gln

Arg Met Ser Ala His Ser Pro Val Pro Pro Pro Leu Ala Pro Leu Ser
 245 250 255
 Ala Pro Ser Gly Thr Ser Ile Ser Ser Ser Glu Gly Cys Tyr Pro
 260 265 270
 Pro Pro His Ala Ala Ala Ser Pro Phe Ala Val Ser Ala Pro Pro Val
 275 280 285
 Met Ala Asp Ser Ala Ala Pro Pro Thr Gly Asn Val Phe Glu Phe His
 290 295 300
 Leu Ala Ala Ala Gly Asp Pro Asn Leu Met Ala Ser Arg Leu Phe Arg
 305 310 315 320
 Ala Arg Arg Ser Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 325 330 335
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 340 345 350
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 355 360 365
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 370 375 380
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Gln Ser
 385 390 395 400
 Pro Pro Pro Pro Pro Leu Leu Pro Ser Phe Asp Ser Ser Cys Ala Gln
 405 410 415
 Ser Ile Pro His Ser Thr Ser Thr Ala Thr Tyr Asp Pro Pro Gly Gly
 420 425 430
 Met Phe Pro Ile Asp Pro Gln Gly His Pro Pro Ser Met Gln Ser Phe
 435 440 445
 Leu Pro Pro Thr Pro Ser Ser Glu Leu Pro Leu Gly Gly Ala Pro Ser
 450 455 460
 Thr His Leu Glu Thr Leu Met Gln Ser Ala Trp Ala Arg His Gly Gly
 465 470 475 480
 Val Val Pro Ala Gln Pro Asp Met Thr Tyr His Glu Ser Leu Leu Ala
 485 490 495
 Met Gln Ala Val Lys Pro Ser Leu Val Leu Leu Pro Phe
 500 505

<210> 14

<211> 2785

<212> DNA

<213> Homo sapiens

<400> 14

cgtgggtggcg	cacgcgtg	gtcccagcta	cttgggaggc	tgaggcagaa	gaatcccttg	60
aaccaggag	gcgagggttg	cagtaagcca	agattgtcac	tgactccag	cctgggagaa	120
agaggaagac	tccatctcaa	aaaaaagaaa	aaagaaattt	catgggttatg	caactcttat	180
ttatgatcag	aaaaatggac	attttgtgat	ttaactctgt	aacatgtttc	atgtagttaa	240
aataataata	aactattaat	catctagctt	gggagagata	ggagaaagac	attactgtca	300
ctagtcaaat	tatatatctt	ttactatcca	ccaaaaatct	cttctgattt	ctgggttagaa	360
ggcatactat	taattgataa	gaaaaataaaa	ctgaaggcct	ctaacatata	acagggtaat	420
aagaatatag	ggaaagttag	ttcaatagtt	ttaattaaag	cacacttctt	acagtataga	480
actagtcggg	cttttatgcc	ttgttttagt	tcttactctt	cctttaactc	ttttctgtt	540
gatgtaattt	acattaatgc	ttaagagtga	actttttaag	tgtgggtaaa	aacgaaataa	600
ttacttacaa	agtttaattc	ttccatttcc	tttgagagag	gaaagttagt	gaaaagcagc	660
tcttatctaa	agcaaagagc	ccacagattg	atttcattgg	ccctggatgt	atttaattgga	720
tttttactat	gcacataatt	tccagaagca	ttgttattta	tttattaatt	ataaatttag	780
tgtaaccatt	tcataagggt	acacagaact	acccagttgt	gcatgtctga	tgtaatttca	840
catatgaatg	tatgaattac	ttgtcttatt	catgttgata	cagcctcagt	ccatggcgca	900
tccgtgtggg	gggaccccaa	catacccaga	atcacagata	tttttcccaa	ctattcatga	960
acgtccagtt	tctttttcac	cacctccac	ctgcccacg	aaagtagcca	tttccagcg	1020
gcgtaagagc	acctccttcc	tggagccca	aactcaccac	ttccaacccc	tgctgaggac	1080
tggtggccaa	agtcttcttc	cacctgggtg	cagcccaact	aactggacac	cagaggccgt	1140
agttatgttg	ggtactacag	ccagtagagt	aactggagag	tcatgtgaga	tacaggcca	1200
tcctatgttt	gaaccatctc	aagtttacag	tgactataga	cctggactag	tacttccaga	1260
agaagctcac	tattttattc	ctcaggaagc	agtgtatgta	gctggggtac	attaccaggc	1320
ccgggtggca	gaacagtatg	agggcattcc	atacaactca	tcagtactgt	caagtcctat	1380
gaaacagata	cctgaacaga	agccagtaca	agggggccct	acttcaagtt	ctgtctttga	1440
tttccatctg	gacaggcttt	cctggtagga	caccttcaga	atctaagatt	agattctgga	1500
ttgggtccgg	gatctcccct	ctctagtatt	tctgcacct	tcagtacaga	tgctacacgt	1560
ttgaaatttc	acctgtctt	tgttcctcat	tctgcgcctg	ctgtgttaac	tcataacaat	1620
gagagcagaa	gcaactgtgt	atgtgaattt	catgttcaca	caccaagctc	ctcttcagga	1680
gaaggagtg	gaattttacc	tcagcgtgtt	taccgaaatc	ggcaggttgc	agtggaacttg	1740
aatcaagaag	aactgcctcc	tcaatcagtt	ggattacatg	gctacttgca	gcctgtgact	1800
gaagaaaagc	ataattacca	tgcccagaa	ttgaccgttt	ctgtggtaga	gcctatcgga	1860
cagaactggc	caataggaag	cccagaatat	tccagtgtat	cctcacaaat	cacttcttca	1920
gaccccgatg	attttcagtc	acctccccct	acagggggag	cagctgcacc	ttttggctct	1980
gacgtctcaa	tgccctttat	ccatctgcct	cagacagtgt	tacaagaatc	cccacttttc	2040
ttctgtttcc	cccaaggaac	cacatctcag	caggtcttaa	ctgcctcatt	ttcttcagga	2100
ggatctgcac	ttcatccaca	ggttatagga	aaacttccac	aattatttta	aactacccta	2160
ctttgcacca	taacatttaa	attttctatt	ccttatttcc	ctgaatcatg	gattttggag	2220
aaatattgtt	taattttatc	agtagagtgt	ccccactctt	ggggggttgt	gaactacata	2280
tatgcattta	aaaacaaaaa	gtgagagaag	ctacctgatt	tacctattat	atgtgaaaac	2340
cagtggaaaa	aacacaaaaa	ctagaatttt	agtcattctt	cacaatcacg	acttctatgc	2400
acgttatattt	caaccagtag	tgaaaatgca	agtgtatgta	atgtatgggt	gaccagcat	2460
tatttaggaa	tacaaatctt	aagtattact	ttcttcctcc	aaacaagttt	ttaaaaata	2520
ggataaattt	tttttctata	aatataaaaa	catggaaaat	agggaaatgct	gtttttgagg	2580
taataattaat	aatacacaga	attttcatta	gtgtcgaagg	atctaaaaag	acaaagtata	2640
tcattgggaat	aaaaaaagat	agaaaaggaa	acagtttagg	aatttgcctt	aacaaatgaa	2700
aatgcctttt	taaaatggca	tcagtcaagc	aagtgtgctg	gcattattat	atgtccaaat	2760
aaaatgctaa	ttcataaaat	taagg				2785

<210> 15
 <211> 2787
 <212> DNA
 <213> Homo sapiens

<400> 15						
cgtgggtggcg	cacgcgtgta	gtcccagcta	cttgggaggc	tgaggcagaa	gaatcccttg	60
aaccaggag	gcgagggttg	cagtaagcca	agattgtcac	tgactccag	cctgggagaa	120
agaggaagac	tccatctcaa	aaaaaagaaa	aaagaaattt	catgggttatg	caactcttat	180
ttatgatcag	aaaaatggac	attttgtgat	ttaactctgt	aacatgtttc	atgtagttaa	240
aataataata	aactattaat	catctagctt	gggagagata	ggagaaagac	attactgtca	300
ctagtcaaat	tatatatctt	ttactatcca	ccaaaaatct	cttctgattt	ctgggttagaa	360

ggcatactat	taattga	gaaaataaaa	ctgaaggcct	ctaacatatc	acagggtaat	420
aagaatatag	ggaaagttag	ttcaatagtt	ttaaattaaag	cacacttctt	acagtataga	480
actagtcggg	cttttatgcc	ttgttttagt	tcttactctt	cctttaactc	tttttctggt	540
gatgtaattt	acattaatgc	ttaagagtga	actttttaag	tgtgggtaaa	aacgaaataa	600
ttacttacaa	agtttaattc	ttccatttcc	tttgagagag	gaaagttatg	gaaaagcagc	660
tcttatctaa	agcaaagagc	ccacagattg	atttcattgg	ccctggatgt	atttaattgga	720
tttttactat	gcacataaatt	tccagaagca	ttgttattta	tttattaatt	ataaatttag	780
tgtaaccatt	tcataggggt	acacagaact	acccagttgt	gcatgtctga	tgtaatttca	840
catatgaatg	tatgaattac	ttgtcttatt	catgttgata	cagcctcagt	ccatggcgca	900
tccgtgtggg	gggaccccaa	catacccaga	atcacagata	tttttcccaa	ctattcatga	960
acgtccagtt	tctttttcac	cacctcccac	ctgcccaccg	aaagtagcca	tttcccagcg	1020
gcgtaagagc	acctccttcc	tggaagccca	aactcaccac	ttccaacccc	tgctgaggac	1080
tgttgggcaa	agtcttcttc	cacctggtgg	cagcccaact	aactggacac	cagaggccgt	1140
agttatgttg	ggtagtacag	ccagtagagt	aactggagag	tcatgtgaga	tacaggcca	1200
tcctatgttt	gaaccatctc	aagtttacag	tgactataga	cctggactag	tactccaga	1260
agaagctcac	tattttattc	ctcaggaagc	agtgtatgta	gctgggtac	attaccaggc	1320
ccgggtggca	gaacagtatg	agggcattcc	atacaactca	tcagtactgt	caagtcctat	1380
gaaacagata	cctgaacaga	agccagtaca	agggggccct	acttcaagtt	ctgtctttga	1440
atttccatct	ggacaggctt	tccctgtagg	acaccttcag	aatctaagat	tagattctgg	1500
attgggtccg	ggatctcccc	tctctagtat	ttctgcacct	atcagtacag	atgctacacg	1560
tttgaaattt	cacctgtct	ttgttctca	ttctgcgcct	gctgtgttaa	ctcataacaa	1620
tgagagcaga	agcaactgtg	tatttgaatt	tcatgttcac	acaccaagct	cctcttcagg	1680
agaaggagggt	ggaattttac	ctcagcgtgt	ttaccgaaat	cggcaggttg	cagtggactt	1740
gaatcaagaa	gaactgcctc	ctcaaatacag	ttggattaca	tggctacttg	cagcctgtga	1800
ctgaagaaaa	gcataattac	catgccccag	aattgaccgt	ttctgtggta	gagcctatcg	1860
gacagaactg	gccaatagga	agcccagaat	attccagtga	ttcctcacia	atcacttctt	1920
cagaccccag	tgattttcag	tcacctcccc	ctacaggggg	agcagctgca	ccttttggct	1980
ctgacgtctc	aatgcccttt	atccatctgc	ctcagacagt	gttaaaagaa	tccccacttt	2040
tcttctgttt	cccccaagga	accacatctc	agcaggtctt	aactgcctca	ttttcttcag	2100
gaggatctgc	acttcatcca	caggttatag	gaaaacttcc	acaattattt	taaactaccc	2160
tactttgcac	cataacattt	aaattttcta	ttccttattt	ccctgaatca	tggatttttg	2220
agaaatattg	tttaatttta	tcagttagagt	ttccccatct	ttgggggggt	gtgaactaga	2280
tatatgcatt	taaaaacaaa	atgtgagaga	agctacctga	tttacctatt	atatgtgaaa	2340
accagtggaa	aaaacacaaa	aactagaatt	ttagtcattc	ttcacaatca	cgacttctat	2400
gcacgttatt	ttcaaccagt	agtgaataatg	caagtgtatg	taatgtatgg	ttgaccagc	2460
attatttagg	aatacaaatc	ttaagtatta	ctttcttcct	ccaaacaagt	ttttaaaaaa	2520
taggataaat	tttttttcta	taaaatataa	aacatggaaa	ataggggaatg	ctgtttttga	2580
ggtaatatata	ataatacaca	gaattttcat	tagtgtcgaa	ggatctaaaa	agacaaaagta	2640
tatcatggga	ataaaaaaag	atagaaaagg	aaacagttta	ggaatttgcc	ttaacaaatg	2700
aaaatgcctt	tttaaaatgg	catcagtcac	gcaagttgct	gtgcattatt	atatgtccaa	2760
ataaaatgct	aattcataaa	attaagg				2787

<210> 16
 <211> 2786
 <212> DNA
 <213> Homo sapiens

<400> 16						
cgtgggtggcg	cacgcgtgta	gtcccagcta	cttggggaggc	tgaggcagaa	gaatcccttg	60
aaccaggag	goggaggttg	cagtaagcca	agattgtcac	tgcaactccag	cctgggagaa	120
agaggaagac	tccatctcaa	aaaaaagaaa	aaagaaattt	catggttatg	caactcttat	180
ttatgatcag	aaaaatggac	attttgtgat	ttaactctgt	aacatgtttc	atgtagtaaa	240
aataataata	aactattaat	catctagctt	gggagagata	ggagaaagac	attactgtca	300
ctagtcaaat	tatatatctt	ttactatcca	ccaaaaatct	cttctgattt	ctggtagaa	360
ggcatactat	taattgataa	gaaaataaaa	ctgaaggcct	ctaacatatc	acagggtaat	420
aagaatatag	ggaaagttag	ttcaatagtt	ttaaattaaag	cacacttctt	acagtataga	480
actagtcggg	cttttatgcc	ttgttttagt	tcttactctt	cctttaactc	tttttctggt	540
gatgtaattt	acattaatgc	ttaagagtga	actttttaag	tgtgggtaaa	aacgaaataa	600
ttacttacaa	agtttaattc	ttccatttcc	tttgagagag	gaaagttatg	gaaaagcagc	660
tcttatctaa	agcaaagagc	ccacagattg	atttcattgg	ccctggatgt	atttaattgga	720

tttttactat	gcacata	tccagaagca	ttgttattta	tttattaatt	atgaatttag	780
tgtaaccatt	tcatagggtt	acacagaact	acccagttgt	gcatgtctga	tgtaatttca	840
catatgaatg	tatgaattac	ttgtcttatt	catgttgata	cagcctcagt	ccatggcgca	900
tccgtgtggg	gggaccccaa	catacccgaga	atcacagata	tttttcccaa	ctattcatga	960
acgtccagtt	tctttttcac	cacctccccc	ctgcccaccg	aaagtagcca	tttcccagcg	1020
gcgtaagagc	acctccttcc	tggaagccca	aactcaccac	ttccaacccc	tgctgaggac	1080
tgttggccaa	agtcttcttc	cacctgggtg	cagcccaact	aactggacac	cagaggccgt	1140
agttatgttg	ggtactacag	ccagtagagt	aactggagag	tcagtgtgaga	tacaggcca	1200
tcctatgttt	gaaccatctc	aagtttacag	tgactataga	cctggactag	tacttccaga	1260
agaagctcac	tattttattc	ctcaggaagc	agtgtatgta	gctgggggtac	attaccaggc	1320
ccgggtggca	gaacagtatg	agggcattcc	atacaactca	tcagtactgt	caagtcctat	1380
gaaacagata	cctgaacaga	agccagtaga	agggggccct	acttcaagtt	ctgtctttga	1440
atttccatct	ggacaggctt	tcctggtagg	apaccttcag	aatctaagat	tagattctgg	1500
attgggtccg	ggatctcccc	tctctagtat	ttctgcacct	atcagtacag	atgctacacg	1560
tttgaaattt	cacctgtctc	ttgttcctca	ttctgcccct	gctgtgttaa	ctcataacaa	1620
tgagagcaga	agcaactgtg	tatttgaatt	tcagtgtcac	acaccaagct	cctcttcagg	1680
agaaggaggt	ggaattttac	ctcagcgtgt	ttaccgaaat	cggcagggtg	cagtggactt	1740
gaatcaagaa	gaactgcctc	ctcaatcagt	tggtattacat	ggctacttgt	agcctgtgac	1800
tgaagaaaag	cataattacc	atgccccaga	attgaccgtt	tctgtggtag	agcctatcgg	1860
acagaactgg	ccaataggaa	gcccagaata	ttccagtgat	tcctcacaac	tcacttcttc	1920
agaccccagt	gattttcagt	cacctcccc	tacaggggga	gcagctgcac	cttttggtct	1980
tgacgtctca	atgcccttta	tccatctgcc	tcagacagt	ttacaagaat	ccccactttt	2040
cttctgtttc	ccccaaaggaa	ccacatctca	gcagggtctta	actgcctcat	tttcttcagg	2100
aggatctgca	cttcatccac	aggttatagg	aaaacttcca	caattatttt	aaactaccct	2160
actttgcacc	ataacattta	aattttctat	tccttatttc	cctgaatcat	ggattttgga	2220
gaaatattgt	ttaattttat	cagtagagtt	tccccatctt	tgggggggtg	tgaactacat	2280
atatgcattt	aaaaacaaaa	tgtgagagaa	gctacctgat	ttacctatta	tatgtgaaaa	2340
ccagtggaaa	aaacacaaaa	actagaattt	tagtcattct	tcacaatcac	gacttctatg	2400
cacgttattt	tcaaccagta	gtgaaaatgc	aagtgtatgt	aatgtatggg	tgaccagca	2460
ttatttagga	atacaaatct	taagtattac	tttcttcctc	caaacaagtt	tttaaaaaat	2520
aggataaatt	ttttttctat	aaaatataaa	acatggaaaa	tagggaaatgc	tgtttttgag	2580
gtaatattaa	taatacacag	aattttcatt	agtgtcgaag	gatctaaaaa	gacaaagtat	2640
atcatgggaa	taaaaaaaaga	tagaaaagga	aacagtttag	gaatttgcct	taacaaatga	2700
aaatgccttt	ttaaaatggc	atcagtcagg	caagttgctg	tgcatatta	tatgtccaaa	2760
taaaatgcta	attcataaaa	ttaagg				2786

<210> 17

<211> 2365

<212> DNA

<213> Sus scrofa

<400> 17

gaattcggca	cgagggtatc	cacaaaaaat	ctcttctgat	ttctaggtag	aagacacact	60
attagttaat	cagaaattaa	ctctggagcc	tctaacagca	cacaacaata	aagcttcaga	120
ataggaaaac	ttggtacagt	ctcattctta	cttaactctg	tttccactga	tgtagtttac	180
attaagtcat	gagtgaacat	attaattgtg	ggtaaaagca	gaataatcac	ttaggaagtt	240
aaatccttcc	atttcccttg	agagaggaga	attgtggaaa	agcagctcct	atctaatagca	300
aagagtccca	cagaataatt	ttattgacct	tggatgtatt	taatggattt	ttttttttta	360
ctgtgcacat	aattttctgaa	agcattgtta	tttctttatt	aattataaat	ttggtgtaac	420
catttctgtg	ggttacacag	agctaccag	ttatgcatgt	ctaattgtaat	ttcacatatg	480
aatgtatgaa	ttacttgtct	tattcatggt	gatacagcct	cagtccatgg	cgcatccgcg	540
tggggggacc	ccaacatacc	cagaatcaca	gatatttttc	ccaactattc	atgaacgtcc	600
agtttctttt	tcaccacctc	ccacctgtcc	accgaagggtg	gctattttcc	agcggcgtaa	660
gagcacctcc	ttcctggaag	cccaaactca	tcacttccaa	cccctgctga	ggactgttgg	720
ccaaaatctt	cttccacctg	gtggctgccc	aactaactgg	acaccagagg	ccgtagtatt	780
gttgggtact	acagccagta	gagtaactgg	agagccatgt	gagatacagg	tccaacctct	840
gtttgagcct	actcaagttt	atgggtgactg	tagacctgga	ctagtagttc	cagaagaagc	900
tcactacttt	attcctcagg	aagcagtgtg	tgtagcaggg	gtacattacc	agacccaaat	960
ggcagaacag	tttgagggtg	ttccatacaa	ctctccagtc	ctgtcaagtc	ctatgaaaca	1020
gatacctgaa	cagaagccag	tgcaaggggg	ccctccttcg	agttctgtct	ttgaatttcc	1080

```

atctggacag gctttcc taggacatct tcagaattta cgattagatt cactactaag 1140
tccaggatct cccctctcta gcatttctac acctatcagt acagatgcta cactgttgaa 1200
atttcaccct gtctttgttc ctcatctctgc acctgctgtg ttaactcata acaatgagag 1260
cagaagcaat tgtgtatttg agtttcatgt tcatactcca agtcctctct caggagaagg 1320
aggagtttta cctcagcgtt tttaccgaaa tcgacagggt gccgtggact tgaatcagga 1380
agaaccacct cctcaatcag ctggattaca cggccgcctg cagcctgtga ctgaagaaca 1440
acataatttc cagccccag aattgaccgt ttctgtggta gagcctactg gacagagctg 1500
gccaatagga agcccagaat attccagtga ttctcacaa atcacttctt cagacccag 1560
tgattttcaa tcacctcccc ctacaggggg aacagctgca ccttttggct ctgacgtctc 1620
attacccttt atccatctgc ctacagacagt gatacaagaa tccccaactt tcttctgttt 1680
cccccaagga accacttctc cgcagatttt atctgcatca ttttcttcag gaggatctgc 1740
actccatcca caggttatag gaaaacttcc acagttctct taaactaccc tgctgtgcac 1800
cattgcattt aagttttctg tccctccttt ccattcatga ggattattgt ttaattttat 1860
tacaatttct ccattttgtg tgtgtgttgg gggggttatt ttaactact tgtatttaag 1920
aacaaaatct gagagcaagc ctgcctgatt tacctataat caaacattg gaaaagaaaa 1980
gaaactagaa ttttagttac ttgtcaaaat cagtgcattc tgttcatatg tagaacgttt 2040
tcaactagaa gtgaaatgta aatgtctgta aagttgactc agaattgatg cttcacacat 2100
actaatgtga agtatttctt tcttcctaga aatgaacgtg gctagttttt taagagaaga 2160
taacattttt ttttcctgta aaatgaaaag aatgggaagt ggagaatcct ttttctaaaa 2220
gtagtattag gagtacatac agaaattaaa agagaggaaa catcttaata agtgaagttg 2280
cctttttaaa atggcatcaa acaaatcaat ttgtattatg tgttcaaata aaatgttaat 2340
ccataaaaaa aaaaaaaaaa aaaaaa 2365

```

<210> 18
 <211> 433
 <212> PRT
 <213> Sus scrofa

<400> 18
 Met Tyr Glu Leu Leu Val Leu Phe Met Leu Ile Gln Pro Gln Ser Met
 1 5 10 15
 Ala His Pro Arg Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile Phe
 20 25 30
 Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Thr
 35 40 45
 Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser Phe
 50 55 60
 Leu Glu Ala Gln Thr His His Phe Gln Pro Leu Leu Arg Thr Val Gly
 65 70 75 80
 Gln Asn Leu Leu Pro Pro Gly Gly Cys Pro Thr Asn Trp Thr Pro Glu
 85 90 95
 Ala Val Val Met Leu Gly Thr Thr Ala Ser Arg Val Thr Gly Glu Pro
 100 105 110
 Cys Glu Ile Gln Val Gln Pro Leu Phe Glu Pro Thr Gln Val Tyr Gly
 115 120 125
 Asp Cys Arg Pro Gly Leu Val Leu Pro Glu Glu Ala His Tyr Phe Ile
 130 135 140
 Pro Gln Glu Ala Val Tyr Val Ala Gly Val His Tyr Gln Thr Gln Met
 145 150 155 160
 Ala Glu Gln Phe Glu Gly Ile Pro Tyr Asn Ser Pro Val Leu Ser Ser

Pro Met Lys Gln Ile Pro Glu Gln Lys Pro Val Gln Gly Gly Pro Pro
 180 185 190

Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val Gly
 195 200 205

His Leu Gln Asn Leu Arg Leu Asp Ser Gly Leu Ser Pro Gly Ser Pro
 210 215 220

Leu Ser Ser Ile Ser Thr Pro Ile Ser Thr Asp Ala Thr Arg Leu Lys
 225 230 235 240

Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr His
 245 250 255

Asn Asn Glu Ser Arg Ser Asn Cys Val Phe Glu Phe His Val His Thr
 260 265 270

Pro Ser Ser Ser Ser Gly Glu Gly Gly Val Leu Pro Gln Arg Ile Tyr
 275 280 285

Arg Asn Arg Gln Val Ala Val Asp Leu Asn Gln Glu Glu Pro Pro Pro
 290 295 300

Gln Ser Ala Gly Leu His Gly Arg Leu Gln Pro Val Thr Glu Glu Gln
 305 310 315 320

His Asn Phe Gln Pro Pro Glu Leu Thr Val Ser Val Val Glu Pro Thr
 325 330 335

Gly Gln Ser Trp Pro Ile Gly Ser Pro Glu Tyr Ser Ser Asp Ser Ser
 340 345 350

Gln Ile Thr Ser Ser Asp Pro Ser Asp Phe Gln Ser Pro Pro Pro Thr
 355 360 365

Gly Gly Thr Ala Ala Pro Phe Gly Ser Asp Val Ser Leu Pro Phe Ile
 370 375 380

His Leu Pro Gln Thr Val Ile Gln Glu Ser Pro Leu Phe Phe Cys Phe
 385 390 395 400

Pro Gln Gly Thr Thr Ser Pro Gln Ile Leu Ser Ala Ser Phe Ser Ser
 405 410 415

Gly Gly Ser Ala Leu His Pro Gln Val Ile Gly Lys Leu Pro Gln Phe
 420 425 430

Ser

<210> 19
 <211> 426
 <212> PRT
 <213> Homo sapiens

<400> 19
 Met Leu Ile Gln Pro Gln Ser Met Ala His Pro Cys Gly Gly Thr Pro

1

5

10

Thr Tyr Pro Glu Ser Gln Ile Phe Phe Pro Thr Ile His Glu Arg Pro
 20 25 30
 Val Ser Phe Ser Pro Pro Pro Thr Cys Pro Pro Lys Val Ala Ile Ser
 35 40 45
 Gln Arg Arg Lys Ser Thr Ser Phe Leu Glu Ala Gln Thr His His Phe
 50 55 60
 Gln Pro Leu Leu Arg Thr Val Gly Gln Ser Leu Leu Pro Pro Gly Gly
 65 70 75 80
 Ser Pro Thr Asn Trp Thr Pro Glu Ala Val Val Met Leu Gly Thr Thr
 85 90 95
 Ala Ser Arg Val Thr Gly Glu Ser Cys Glu Ile Gln Val His Pro Met
 100 105 110
 Phe Glu Pro Ser Gln Val Tyr Ser Asp Tyr Arg Pro Gly Leu Val Leu
 115 120 125
 Pro Glu Glu Ala His Tyr Phe Ile Pro Gln Glu Ala Val Tyr Val Ala
 130 135 140
 Gly Val His Tyr Gln Ala Arg Val Ala Glu Gln Tyr Glu Gly Ile Pro
 145 150 155 160
 Tyr Asn Ser Ser Val Leu Ser Ser Pro Met Lys Gln Ile Pro Glu Gln
 165 170 175
 Lys Pro Val Gln Gly Gly Pro Thr Ser Ser Ser Val Phe Glu Phe Pro
 180 185 190
 Ser Gly Gln Ala Phe Leu Val Gly His Leu Gln Asn Leu Arg Leu Asp
 195 200 205
 Ser Gly Leu Gly Pro Gly Ser Pro Leu Ser Ser Ile Ser Ala Pro Ile
 210 215 220
 Ser Thr Asp Ala Thr Arg Leu Lys Phe His Pro Val Phe Val Pro His
 225 230 235 240
 Ser Ala Pro Ala Val Leu Thr His Asn Asn Glu Ser Arg Ser Asn Cys
 245 250 255
 Val Phe Glu Phe His Val His Thr Pro Ser Ser Ser Ser Gly Glu Gly
 260 265 270
 Gly Gly Ile Leu Pro Gln Arg Val Tyr Arg Asn Arg Gln Val Ala Val
 275 280 285
 Asp Leu Asn Gln Glu Glu Leu Pro Pro Gln Ser Val Gly Leu His Gly
 290 295 300
 Tyr Leu Gln Pro Val Thr Glu Glu Lys His Asn Tyr His Ala Pro Glu
 305 310 315 320
 Leu Thr Val Ser Val Val Glu Pro Ile Gly Gln Asn Trp Pro Ile Gly
 325 330 335

Ser Pro Glu Tyr Ser Ser Asp Ser Ser Gln Ile Thr Ser Ser Asp Pro
 340 345 350

Ser Asp Phe Gln Ser Pro Pro Pro Thr Gly Gly Ala Ala Ala Pro Phe
 355 360 365

Gly Ser Asp Val Ser Met Pro Phe Ile His Leu Pro Gln Thr Val Leu
 370 375 380

Gln Glu Ser Pro Leu Phe Phe Cys Phe Pro Gln Gly Thr Thr Ser Gln
 385 390 395 400

Gln Val Leu Thr Ala Ser Phe Ser Ser Gly Gly Ser Ala Leu His Pro
 405 410 415

Gln Val Ile Gly Lys Leu Pro Gln Leu Phe
 420 425

<210> 20
 <211> 419
 <212> PRT
 <213> Homo sapiens

<400> 20
 Met Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile
 1 5 10 15

Phe Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Pro
 20 25 30

Thr Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser
 35 40 45

Phe Leu Glu Ala Gln Thr His His Phe Gln Pro Leu Leu Arg Thr Val
 50 55 60

Gly Gln Ser Leu Leu Pro Pro Gly Gly Ser Pro Thr Asn Trp Thr Pro
 65 70 75 80

Glu Ala Val Val Met Leu Gly Thr Thr Ala Ser Arg Val Thr Gly Glu
 85 90 95

Ser Cys Glu Ile Gln Val His Pro Met Phe Glu Pro Ser Gln Val Tyr
 100 105 110

Ser Asp Tyr Arg Pro Gly Leu Val Leu Pro Glu Glu Ala His Tyr Phe
 115 120 125

Ile Pro Gln Glu Ala Val Tyr Val Ala Gly Val His Tyr Gln Ala Arg
 130 135 140

Val Ala Glu Gln Tyr Glu Gly Ile Pro Tyr Asn Ser Ser Val Leu Ser
 145 150 155 160

Ser Pro Met Lys Gln Ile Pro Glu Gln Lys Pro Val Gln Gly Gly Pro
 165 170 175

Thr Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val

Gly His Leu Gln Asn Leu Arg Leu Asp Ser Gly Leu Gly Pro Gly Ser
 195 200 205
 Pro Leu Ser Ser Ile Ser Ala Pro Ile Ser Thr Asp Ala Thr Arg Leu
 210 215 220
 Lys Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr
 225 230 235 240
 His Asn Asn Glu Ser Arg Ser Asn Cys Val Phe Glu Phe His Val His
 245 250 255
 Thr Pro Ser Ser Ser Ser Gly Glu Gly Gly Gly Ile Leu Pro Gln Arg
 260 265 270
 Val Tyr Arg Asn Arg Gln Val Ala Val Asp Leu Asn Gln Glu Glu Leu
 275 280 285
 Pro Pro Gln Ser Val Gly Leu His Gly Tyr Leu Gln Pro Val Thr Glu
 290 295 300
 Glu Lys His Asn Tyr His Ala Pro Glu Leu Thr Val Ser Val Val Glu
 305 310 315 320
 Pro Ile Gly Gln Asn Trp Pro Ile Gly Ser Pro Glu Tyr Ser Ser Asp
 325 330 335
 Ser Ser Gln Ile Thr Ser Ser Asp Pro Ser Asp Phe Gln Ser Pro Pro
 340 345 350
 Pro Thr Gly Gly Ala Ala Ala Pro Phe Gly Ser Asp Val Ser Met Pro
 355 360 365
 Phe Ile His Leu Pro Gln Thr Val Leu Gln Glu Ser Pro Leu Phe Phe
 370 375 380
 Cys Phe Pro Gln Gly Thr Thr Ser Gln Gln Val Leu Thr Ala Ser Phe
 385 390 395 400
 Ser Ser Gly Gly Ser Ala Leu His Pro Gln Val Ile Gly Lys Leu Pro
 405 410 415
 Gln Leu Phe

<210> 21
 <211> 425
 <212> PRT
 <213> Mus musculus

<400> 21
 Met Leu Ile Gln Pro Gln Ser Met Ala His Pro Cys Gly Gly Thr Pro
 1 5 10 15
 Thr Tyr Pro Glu Ser Gln Ile Phe Phe Pro Thr Ile His Glu Arg Pro
 20 25 30
 Val Ser Phe Ser Pro Pro Pro Thr Cys Pro Pro Lys Val Ala Ile Ser

Gln Arg Arg Lys Ser Thr Ser Phe Leu Glu Ala Gln Thr Arg His Phe
 50 55 60
 Gln Pro Leu Leu Arg Thr Val Gly Gln Asn His Leu Pro Pro Gly Ser
 65 70 75 80
 Ser Pro Thr Asn Trp Thr Pro Glu Ala Ile Val Met Leu Gly Ala Thr
 85 90 95
 Ala Asn Arg Val Asn Arg Glu Leu Cys Glu Met Gln Val Gln Pro Val
 100 105 110
 Phe Glu Pro Thr Gln Ile Tyr Ser Asp Tyr Arg Pro Gly Leu Val Leu
 115 120 125
 Ala Glu Glu Ala His Tyr Phe Ile Pro Gln Glu Thr Val Tyr Leu Ala
 130 135 140
 Gly Val His Tyr Gln Ala Gln Val Ala Gly Gln Tyr Glu Gly Ile Ser
 145 150 155 160
 Tyr Asn Ser Pro Val Leu Ser Ser Pro Met Lys Gln Ile Ser Glu Gln
 165 170 175
 Lys Pro Val Pro Gly Gly Pro Ala Ser Ser Ser Val Phe Glu Phe Pro
 180 185 190
 Ser Gly Gln Ala Phe Leu Val Gly His Leu Gln Asn Leu Arg Leu Asp
 195 200 205
 Ser Gly Pro Ser Pro Ala Ser Pro Leu Ser Ser Ile Ser Ala Pro Asn
 210 215 220
 Ser Thr Asp Ala Thr His Leu Lys Phe His Pro Val Phe Val Pro His
 225 230 235 240
 Ser Ala Pro Ala Val Leu Thr Asn Ser Asn Glu Asn Arg Ser Asn Cys
 245 250 255
 Val Phe Glu Phe His Ala Gln Thr Pro Ser Ser Ser Gly Glu Gly Gly
 260 265 270
 Gly Ile Leu Pro Gln Arg Val Tyr Arg Asn Arg Gln Val Ala Val Asp
 275 280 285
 Ser Asn Gln Glu Glu Leu Ser Pro Gln Ser Val Gly Leu His Cys His
 290 295 300
 Leu Gln Pro Val Thr Glu Glu Gln Arg Asn Asn His Ala Pro Glu Leu
 305 310 315 320
 Thr Ile Ser Val Val Glu Pro Met Gly Gln Ile Trp Pro Ile Gly Ser
 325 330 335
 Pro Glu Tyr Ser Ser Asp Ser Ser Gln Ile Thr Ser Ser Asp Leu Ser
 340 345 350
 Asp Phe Gln Ser Pro Pro Pro Thr Gly Gly Thr Ala Ala Pro Phe Gly
 355 360 365

Ser Asp Val Ser Leu Pro Phe Ile Arg Leu Pro Gln Thr Val Leu Gln
 370 375 380

Glu Ser Pro Leu Phe Phe Cys Phe Pro Gln Gly Thr Thr Ser Gln Gln
 385 390 395 400

Val Leu Ser Ala Ser Tyr Ser Ser Gly Gly Ser Thr Leu His Pro Gln
 405 410 415

Val Ile Gly Lys Leu Ser Gln Phe Phe
 420 425

<210> 22
 <211> 418
 <212> PRT
 <213> Mus musculus

<400> 22

Met Ala His Pro Cys Gly Gly Thr Pro Thr Tyr Pro Glu Ser Gln Ile
 1 5 10 15

Phe Phe Pro Thr Ile His Glu Arg Pro Val Ser Phe Ser Pro Pro Pro
 20 25 30

Thr Cys Pro Pro Lys Val Ala Ile Ser Gln Arg Arg Lys Ser Thr Ser
 35 40 45

Phe Leu Glu Ala Gln Thr Arg His Phe Gln Pro Leu Leu Arg Thr Val
 50 55 60

Gly Gln Asn His Leu Pro Pro Gly Ser Ser Pro Thr Asn Trp Thr Pro
 65 70 75 80

Glu Ala Ile Val Met Leu Gly Ala Thr Ala Asn Arg Val Asn Arg Glu
 85 90 95

Leu Cys Glu Met Gln Val Gln Pro Val Phe Glu Pro Thr Gln Ile Tyr
 100 105 110

Ser Asp Tyr Arg Pro Gly Leu Val Leu Ala Glu Glu Ala His Tyr Phe
 115 120 125

Ile Pro Gln Glu Thr Val Tyr Leu Ala Gly Val His Tyr Gln Ala Gln
 130 135 140

Val Ala Gly Gln Tyr Glu Gly Ile Ser Tyr Asn Ser Pro Val Leu Ser
 145 150 155 160

Ser Pro Met Lys Gln Ile Ser Glu Gln Lys Pro Val Pro Gly Gly Pro
 165 170 175

Ala Ser Ser Ser Val Phe Glu Phe Pro Ser Gly Gln Ala Phe Leu Val
 180 185 190

Gly His Leu Gln Asn Leu Arg Leu Asp Ser Gly Pro Ser Pro Ala Ser
 195 200 205

Pro Leu Ser Ser Ile Ser Ala Pro Asn Ser Thr Asp Ala Thr His Leu

Lys Phe His Pro Val Phe Val Pro His Ser Ala Pro Ala Val Leu Thr
 225 230 235 240
 Asn Ser Asn Glu Asn Arg Ser Asn Cys Val Phe Glu Phe His Ala Gln
 245 250 255
 Thr Pro Ser Ser Ser Gly Glu Gly Gly Gly Ile Leu Pro Gln Arg Val
 260 265 270
 Tyr Arg Asn Arg Gln Val Ala Val Asp Ser Asn Gln Glu Glu Leu Ser
 275 280 285
 Pro Gln Ser Val Gly Leu His Cys His Leu Gln Pro Val Thr Glu Glu
 290 295 300
 Gln Arg Asn Asn His Ala Pro Glu Leu Thr Ile Ser Val Val Glu Pro
 305 310 315 320
 Met Gly Gln Ile Trp Pro Ile Gly Ser Pro Glu Tyr Ser Ser Asp Ser
 325 330 335
 Ser Gln Ile Thr Ser Ser Asp Leu Ser Asp Phe Gln Ser Pro Pro Pro
 340 345 350
 Thr Gly Gly Thr Ala Ala Pro Phe Gly Ser Asp Val Ser Leu Pro Phe
 355 360 365
 Ile Arg Leu Pro Gln Thr Val Leu Gln Glu Ser Pro Leu Phe Phe Cys
 370 375 380
 Phe Pro Gln Gly Thr Thr Ser Gln Gln Val Leu Ser Ala Ser Tyr Ser
 385 390 395 400
 Ser Gly Gly Ser Thr Leu His Pro Gln Val Ile Gly Lys Leu Ser Gln
 405 410 415
 Phe Phe

<210> 23
 <211> 426
 <212> PRT
 <213> Ratus ratus

<400> 23
 Met Leu Ile Gln Pro Gln Ser Val Ala His Pro Cys Gly Gly Thr Pro
 1 5 10 15
 Thr Tyr Pro Glu Ser Gln Ile Phe Phe Pro Thr Ile His Glu Arg Pro
 20 25 30
 Val Ser Phe Ser Pro Pro Pro Thr Cys Pro Pro Lys Val Ala Ile Ser
 35 40 45
 Gln Arg Arg Lys Ser Thr Ser Phe Leu Glu Ala Gln Thr Arg His Phe
 50 55 60
 Gln Pro Leu Leu Arg Thr Val Gly Gln Asn His Leu Pro Pro Gly Gly

Ser Pro Thr Asn Trp Thr Pro Glu Ala Ile Val Met Leu Gly Thr Thr
 85 90 95
 Ala Asn Arg Val Asn Arg Glu Leu Cys Glu Met Gln Val Gln Pro Val
 100 105 110
 Phe Glu Thr Thr Gln Ile Tyr Ser Asp Tyr Arg Pro Gly Leu Val Leu
 115 120 125
 Ala Glu Glu Ala His Tyr Phe Ile Pro Gln Glu Thr Val Tyr Leu Ala
 130 135 140
 Gly Val His Tyr Gln Ala His Ala Ala Gly Gln Tyr Glu Gly Ile Ser
 145 150 155 160
 Tyr Asn Ser Pro Val Leu Ser Ser Pro Met Lys Gln Ile Thr Glu Gln
 165 170 175
 Lys Pro Val Pro Gly Cys Pro Ala Ser Ser Ser Val Phe Glu Phe Pro
 180 185 190
 Ser Gly Gln Ala Phe Leu Val Gly His Leu Gln Asn Leu Arg Leu Asp
 195 200 205
 Ser Gly Pro Ser Pro Ala Ser Pro Leu Ser Ser Ile Ser Ala Pro Asn
 210 215 220
 Ser Thr Asp Ala Thr His Leu Lys Phe His Pro Val Phe Val Pro His
 225 230 235 240
 Ser Ala Pro Ala Val Leu Thr His Ser Asn Glu Asn Arg Ser Asn Cys
 245 250 255
 Val Phe Glu Phe His Ala Gln Thr Pro Ser Ser Ser Ser Gly Glu Gly
 260 265 270
 Gly Gly Ile Leu Pro Gln Arg Val Tyr Arg Asn Arg Gln Val Ala Val
 275 280 285
 Asp Ser Ser Gln Glu Glu Leu Ser Pro Gln Ser Val Gly Leu His Cys
 290 295 300
 His Leu Gln Pro Val Thr Glu Glu Gln Arg Asn Asn His Thr Pro Glu
 305 310 315 320
 Leu Thr Ile Ser Val Val Glu Pro Met Gly Gln Asn Trp Pro Val Gly
 325 330 335
 Ser Pro Glu Tyr Ser Ser Asp Ser Ser Gln Ile Thr Ser Ser Asp Ile
 340 345 350
 Ser Asp Phe Gln Ser Pro Pro Pro Thr Gly Gly Thr Ala Ala Pro Phe
 355 360 365
 Gly Ser Asp Val Ser Leu Pro Tyr Ile Arg Leu Pro Gln Thr Val Leu
 370 375 380
 Gln Glu Ser Pro Leu Phe Phe Cys Phe Pro Gln Gly Thr Thr Ser Gln
 385 390 395 400

Gln Val Leu Ser Ala Ser Tyr Ser Ser Gly Gly Ser Ala Leu His Pro
405 410 415

Gln Val Ile Gly Lys Leu Ser Gln Phe Phe
420 425

<210> 24

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Forward amplification Primer

<400> 24

ttccagaagc attggtattt attt

24

<210> 25

<211> 19

<212> DNA

<213> Artificial

<220>

<223> Reverse Replication Primer

<400> 25

cccccttgta ctggcttct

19

<210> 26

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Forward Replication Primer

<400> 26

caccagaggc cgtagttatg ttg

23

<210> 27

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Reverse replication primer.

<400> 27

ttgaggaggc agttcttctt gatt

24

<210> 28

<211> 23

<212> DNA

<213> Artificial

<220>
<223> Forward replication primer.

<400> 28
gcgcctgctg tgttaactca taa 23

<210> 29
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Reverse replication primer.

<400> 29
ccaaagatgg ggaaactcta ctga 24

<210> 30
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Forward replication primer.

<400> 30
accatcacct aaggagacag acc 23

<210> 31
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Reverse replication primer.

<400> 31
tgcaacaaat gtaccactct gg 22

<210> 32
<211> 16
<212> DNA
<213> Artificial

<220>
<223> Forward replication primer.

<400> 32
agctcctaaa tccccg 16

<210> 33
<211> 21
<212> DNA
<213> Artificial

<220>

WO 2004/043999

T/CA2003/001716

<223> Reverse replication primer.

<400> 33

gccatgtcta taaataccct g

21

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.